LIBRARY

AUG 3 1935

Vol. XVIII, No. 4 *PSYCHOLOGICAL REVIEW PUBLICATIONS* August, 1935

18

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA

WALTER S. HUNTER, CLARK UNIVERSITY (*Psychol. Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Psychol. Rev.*)
JOSEPH PETERSON, GEORGE PEABODY COLLEGE (*Psychol. Monog.*)
JOHN A. McGEOCH, WESLEYAN UNIVERSITY (*Psychol. Bull.*)

CONTENTS

Binocular Summation in Scotopic Vision: D. SHAAD 391

The Relationship between Time Spent in the Culs-de-sac of a Stylus Maze and Speed of Elimination: H. N. PETERS AND J. A. McGEOCH 414

The Development of the Reactive Process from Early Childhood to Maturity: F. L. GOODENOUGH 431

An Atmosphere Effect in Formal Syllogistic Reasoning: R. S. WOODWORTH AND S. B. SELLS 451

An Experimental Study of the Pairing of Certain Auditory and Visual Stimuli: J. T. COWLES 461

A Preliminary Study of Bisection of Lines: L. CHIDESTER 470

The Frois-Wittmann Pictures of Facial Expression: W. S. HULIN AND D. KATZ 482

The After-Effect of the Perception of Curved Lines: J. F. BALES AND G. L. FOLLANSBEE 499

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*J. Exper. Psych.*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index and Abstracts*)
HENRY T. MOORE, SKIDMORE COLLEGE (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Review*)
JOSEPH PETERSON, GEORGE PEABODY COLLEGE (*Monographs*)
JOHN A. McGEOCH, WESLEYAN UNIVERSITY (*Bulletin*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 900 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consist of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, April, July, October, January, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: $5.50 (Foreign, $5.75). **Index:** $4.00 per volume.
Journal: $7.00 (Foreign, $7.25). **Monographs:** $6.00 per volume (Foreign, $6.30).
Bulletin: $6.00 (Foreign, $6.25). **Abstracts:** $6.00 (Foreign, $6.25).
Abnormal and Social: $5.00 (Foreign, $5.25). Single copies $1.50.
Current numbers: **Journal,** $1.25; **Review,** $1.00; **Abstracts,** 75c; **Bulletin,** 60c.

COMBINATION RATES

Review and Bulletin: $10.00 (Foreign, $10.50).
Review and J. Exp.: $11.00 (Foreign, $11.50).
Bulletin and J. Exp.: $12.00 (Foreign, $12.50).
Review, Bulletin, and J. Exp.: $16.00 (Foreign, $16.75).
Review, Bulletin, J. Exp., and Index: $19.00 (Foreign, $19.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, NEW JERSEY

Journal of
Experimental Psychology

VOL. XVIII, No. 4 AUGUST, 1935

BINOCULAR SUMMATION IN SCOTOPIC VISION *

BY DOROTHY SHAAD

Bryn Mawr College

Investigations of different aspects of visual perception frequently raise the question of whether or not stimuli appear brighter when seen with both eyes than when they are seen with only one eye. Summation, defined as an additive effect occurring within the central nervous system which is mediated by binocular perception and characterized by an increase in the apparent brightness of a given field, is a factor which may influence measurements of visual sensitivity under conditions of dark adaptation. The semi-decussation of the optic tracts, homolateral parts of the retinae being represented on corresponding cortical hemispheres, is a fact which has led to the formation of various theories concerning the interaction of the two retinal fields, which, after neural elaboration, are united to form a single perceptual field.

Although the problem of summation has been primary in a number of studies, there has been such variation in the methods of experiment, combined many times with insufficient

* This study is condensed from a dissertation prepared under the direction of Dr. H. Helson and accepted by Bryn Mawr College in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Copies of the dissertation are on file at the Bryn Mawr College Library. The writer is indebted to Dr. Helson for generous assistance in the experimental work and constructive criticism in the preparation of the thesis.

precaution in the making of measurements, that authorities have failed to agree as to whether or not summation plays a significant rôle in binocular vision. A survey of the literature, since Piper's earliest comments on the presence of summation in scotopic vision (**28**), reveals over thirty years of varied investigation, including both clinical studies and laboratory experiments related to the question of summation. The consensus of opinion, as evident in the older studies, and as supported by such authority as Nagel (**19**) is that summation of brightness exists for scotopic vision. Fröbes (**14**) accepts Piper's statement that summation occurs under conditions of dark adaptation; this view is supported, from experimental evidence, by a number of investigators, among them being von Behr (**6**), DeSilva and Bartley (**9**), Feilchenfeld and Loeser (**11**), Lohmann (**24**), Messmer (**25**), Roelofs and Zeemann (**31**) and Stargardt (**33**). Others, such as Abney and Watson (**1**), Best (**7**), Downey (**10**), Graham (**15**), Müller (**26**) and Sherrington (**32**) have failed to obtain results which they could interpret as indicating the presence of summation in binocular observation.

Although foveal summation has frequently been doubted, Cook (**5**) in a careful study has established the presence of binocular summation for central vision in the dark-adapted eye. The peripheral regions, unlike the foveal regions which have double cortical representation since fibers from each macular region go to both hemispheres, are supplied with nerve fibers whose elaborations carry them to either one hemisphere or the other. Right halves of each peripheral retina are represented in the right cortex; left halves of each retina are correspondingly represented in the left cortex (Peter, **27**). If central summation exists for binocular vision, the stimulation of retinal areas which are elaborated in the same central region should be more effective than the stimulation of retinal areas not elaborated in the same hemisphere. It should be possible to demonstrate the existence of binocular summation in scotopic vision through a comparison of monocular and binocular measurements of liminal light sensitivity of the peripheral retina. On account of the lack of agreement among previous

studies, in order to clarify the evidence either for or against binocular summation, the present study was undertaken.

Experimental Precautions in the Measurement of Visual Sensitivity

The method adopted for measurement of liminal sensitivity to light under conditions of dark adaptation cannot be determined without consideration of a number of factors, which, if uncontrolled, may obscure the true values. The rapid initial increase in sensitivity which occurs in early dark adaptation makes it difficult to obtain accurate measurements of the threshold until sensitivity has reached a fairly stable level. Though dark adaptation has been shown to continue for as long as twenty-four hours (Achmatov, **2**), there is a relatively small change between the level attained after one half-hour in the dark and the next hour's adaptation. Largely because of the difficulty involved in making accurate measurements at specified intervals during the course of adaptation, it is convenient, for any study not concerned primarily with the course of adaptation, to stabilize the factor of rapidly changing sensitivity to light by the use of a definite period of adaptation to precede the measurements made for theoretical use. A half-hour period in the dark serves to restore visual purple to nearly its maximum concentration so that no influence of previous light adaptation on liminal sensitivity can be discerned after this period in the dark.

Since the sensitivity of the dark-adapted eye increases as the distance from the fovea is increased, reaching a maximum 10 to 15 degrees toward the periphery, fixation must be controlled in order to test a given retinal region. Foveal fixation is readily maintained by the use of a cross of red light, since form perception (visual acuity) as well as sensitivity to red is greatest at the fovea, for scotopic vision.

The peripheral retina, with its high sensitivity under conditions of scotopic vision, has a rapid rate of physiological adaptation.[1] Since a light of minimal intensity exposed to

[1] 'Physiological' adaptation refers to changes in the end-organ which lessen its responsiveness; it is distinguished from the physico-chemical changes of dark adaptation.

one retinal region will presently disappear completely, and since liminal measurements demand controlled fixation which may in itself cause fatigue, the use of sufficient recovery, or rest periods during the taking of measurements is necessary. A method in which the exposure of the stimulus is at a minimum, requiring only a short period of steady fixation, is preferable to one requiring long fixations.

Since the size of the pupil for any given individual becomes relatively constant after a few minutes in the dark, and since the intensities at which the usual studies of dark adaptation are made are not great enough to cause any marked pupillary changes,[2] measurements valid for purposes of comparison may be made without the use of artificial pupils.

Spontaneous retinal interference, in the form of the "streaming phenomenon," entoptic colors, *etc.* may be a constant source of inaccuracy in the making of liminal judgments of visibility. The autokinetic sensation, which may interfere with fixation, is likewise detrimental to good observation. These factors necessitate the use of short fixations and frequent recovery periods.

Probably chief among the conditions which should be controlled by instruction in reliable work in visual sensitivity is that of the criteria of judgment to be adopted. The criterion of "just visible light" is not sufficient without further definition, for, as pointed out by Helson and Fehrer (**20**), a limen for mere detection of light as distinguished from its background of darkness is lower than one for recognition of geometrical form. In order to reduce the variation resulting from the practise of judging without regard for qualitative differences in the appearance of the stimuli, it is necessary to use careful instructions to establish definitely the criteria of observation to be used throughout an experiment by the different observers.

Even when the criteria of judgment have been established

[2] Blanchard (**8**, p. 96) gives a table of pupillary size as influenced by changes in illumination. The smallest change in illumination included in the table is from 0.0 ml. to .00015 ml., causing a pupillary change, both eyes open, from 7.4 to 7.15 mm in diameter. Threshold sensitivity under dark adaptation is well below the intensity .00015 ml.

by instruction, an opportunity for the repetition of the same test conditions with the same observers is necessary in order to rule out the effects of unfamiliarity with the experimental procedure. Practise should result in a decrease in the variability of observations taken from one observer. Obviously, this stabilization of individual reports, which is not an habitual, mechanical result of practise, but the effect of a better understanding of the experimental situation, will not abolish the individual variation characteristic of all measurement of liminal sensitivity.

For work in light perception, any instrument which provides measurable and standardized continuous variation in illumination, from low to high intensities, is efficient. Since the visual angle subtended by the stimulus surface, a function of the size of the stimulus and its distance from the eye, determines to a large extent the absolute intensity effective under conditions of dark adaptation, this factor should be carefully regulated.

Statistical treatment of data requires a large number of observations, necessitating either the testing of many individuals, or repeated tests of the same persons. Results obtained from single tests of only a few persons, especially if the measurements are taken during the course of early dark adaptation, are almost certain to be unreliable.

The method usually employed in investigations of liminal light sensitivity is that of gradual increase (or decrease) in stimulus intensity until light first appears (or disappears). A method which retains the advantages of simple, continuous variation, and adds the advantage of brief, controlled exposure times, which are desirable for work under conditions of dark adaptation, consists of discrete stimulus values, arranged in series having physically equivalent intervals, and presented in ascending or descending sequence.[3] For vision, the series may be confined to ascending steps, since after-images may result from changes from supraliminal to subliminal intensities in descending series. Discrete stimuli, in series of varied

[3] This method, often employed in auditory experiments, has been used in unpublished studies in vision at Bryn Mawr College.

length, with different starting points below the threshold, can be used for short exposures. The observer has merely to report the presence or absence of sensation.

Studies in which summation has been considered frequently reveal failure to control one or more of these conditions which may easily invalidate the procedure. In many cases, too few data are secured to assure statistical reliability. In other studies, conclusions with regard to the presence or absence of summation have not always been in keeping with the experimental data presented. The data of Sherrington (**32**) for instance show a slight excess in brightness of the binocular image over the two equal monocular components, but central summation is not postulated.

Conditions of Experiment

The test surface of a Nagel Adaptometer, a standard instrument for measuring changes in light sensitivity,[4] was provided with special screens which contained circular perforations which served as the light stimuli. The perforations were designed to subtend, at a distance of 40 cm from the observer's eyes, angles of 1, 2, 4, or 8 degrees. The center of each spot was 10 degrees from the point of fixation. The size and number of test spots will be indicated for each limen.

The fixation point consisted of a red cross, having an area of 9 sq mm; each arm of the cross was 2.5 mm in length and 1 mm in width. The cross, cut in the side of a flattened, light-proof cardboard cylinder, was covered with red gelatin and illuminated indirectly by reflection from the white inner surface of the cylinder. The light source was a flashlight bulb used with 2.5 volts. The cross, in the same vertical plane as the test spots, appeared directly before the eyes of the observer, which were kept at a fixed distance (40 cm) by means of a stationary headrest. The test spots were at a level slightly above that of the fixation light, the lowest spots used being 2.89 degrees above a horizontal passing through the point of fixation. The arrangement of stimuli, 10 degrees from the fixation point and somewhat above it, assured complete freedom from any disturbance by conditions near the slind spot. The projected blind spots of each observer as mapped against the actual experimental field were well below and to the side of the stimulus spot.

The central neurological connections which form the theoretical basis for a comparison of monocular and binocular measurements of liminal light sensitivity are shown schematically in Fig. I. Left retinal fibers are indicated by solid black; right retinal fibers are shown in outline, in order to indicate the semi-decussation of the fibers at the chiasm. Homolateral cortical representation in the occipital regions, as well as

[4] For a detailed description of this instrument, Helmholtz (**19**, pp. 315 ff.) may be consulted. The current supplying the light source was maintained at 105 volts by the use of variable resistance. By means of the Macbeth Illuminometer, a correction factor for converting the scale values of the adaptometer into absolute intensities in millilamberts was calculated. This factor was: 4.78×10^{-9}.

the macular fibers passing to both hemispheres from each retina, are omitted from the diagram. Figure 1 shows, however, the arrangement of test spots, one on either side of the fixation point and of the removable screens (*a*, *b*, and *c*) which were provided in order to vary the conditions of experiment. The screens, supported by a stand before the observer's headrest, were cut from cardboard and were designed in such a way that

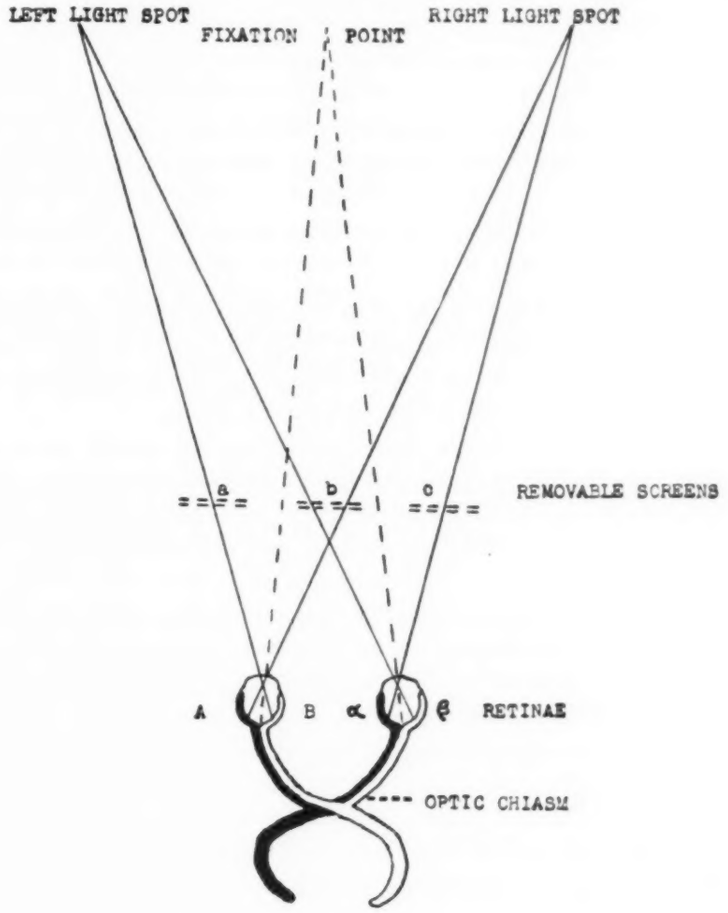


FIG. 1. Experimental schema and diagrammatic representation of part of the visual tract, showing the semi-decussation of the fibers of the optic nerves at the chiasm.

it was not necessary to interfere with binocular foveal fixation of the red cross. One or more screens could be used at the same time. Screen *a* prevented stimulation of the left eye by the left light spot; screen *b* prevented stimulation of the temporal regions of either eye; screen *c* prevented stimulation of the right eye by the right light spot. The retinal regions, indicated by *A*, *B*, α and β, in Fig. 1, will be referred to by these letters, in order to identify the retinal points stimulated under the conditions of experiment. Various combinations of retinal points simultaneously stimulated will be referred to by the use of two or more symbols.

The following limens for controlled retinal stimulation were measured:

I. Limens using a single 2-degree light spot.
   1. *A* Monocular; left temporal retina.
   2. α " right nasal retina.

3. *Aα* Binocular; right nasal and left temporal corresponding points.
4. *B* Monocular; left nasal retina.
5. *β* " right temporal retina.
6. *Bβ* Binocular; right temporal and left nasal corresponding points.

II. Limens using two 2-degree light spots, one on either side of the fixation point.

7. *Aβ* Temporal non-corresponding points of the two eyes.
8. *Bα* Nasal " " " " " " "
9. *AB* Monocular non-corresponding points (left eye).
10. *αβ* " " " " (right eye).
11. *AαBβ* Binocular corresponding points of the two eyes.
12. *ABβ* Monocular non-corresponding points of left eye; right temporal retina.
13. *Aαβ* Monocular non-corresponding points of right eye; left temporal retina.

Limens 9 and 10, in which monocular non-corresponding[5] points were tested, constitute an exception in that for these tests it was necessary to employ a special screen, set close to the observer's eye, which obscured the stimuli from the unused eye, but at the same time prevented binocular fixation of the red cross.

Each observation period, during which it was possible usually to secure data for two limens, was preceded by a 25-minute period of dark adaptation. A peripheral region 10 degrees from the fovea was chosen for making the liminal measurements, since this area has been shown (Wölfflin, **34**) to be of maximum sensitivity under conditions of dark adaptation.

The normal pupil was used since the stimulus intensity, which was never greater than .00001336 ml., was insufficient to cause significant variation in pupillary size (see note 2). The brightness of the fixation light, a constant factor, was too low to be measured by the Macbeth Illuminometer.

For the following experiments, five observers, including the experimenter, were employed: V. Balough (B), graduate student in psychology; E. Fehrer (F), graduate student in psychology; Dr. H. Helson (H), professor of psychology; D. Shaad (S), experimenter; Dr. W. Turner (T), instructor in psychology. Each observer had had previous experience in visual experiments.

Procedure and Results

The first procedure used in obtaining limens consisted of gradual, continuous increase in the intensity of illumination, controlled by the observer himself, until light (not form) was reported as 'just visible.' The threshold values, determined by averaging the judgments of just visible light in 50 trials, showed wide dispersion. This procedure, frequently employed in studies of liminal light sensitivity, proved unsatisfactory because, though variations were usually in the same direction for the different observers, differences were not statistically reliable. Adaptation to subliminal stimulation, combined with long fixations and irregularity in the rate of increase of stimulus intensity were considered major factors responsible for the instability of the results obtained by this method.

In the second procedure, short exposures of discrete stimulus intensities were used

[5] 'Non-corresponding' is used here with reference to retinal regions centrally projected on different hemispheres, regardless of whether the retinal areas stimulated are situated on either side of the same retina, or one on each retina.

in ascending series; the observer reported the point at which light first became visible. The short exposures gave less time for adaptation and prevented loss of fixation or eye movements likely to accompany long fixations. There was no variation in rate of change of intensity among the different observers by this method, since the observer was required merely to fixate at the signal 'Ready' and report the presence or absence of the stimulus following the 'Now' signal given by the experimenter. The stimuli were exposed for approximately 2.5 seconds by means of a movable screen placed in front of the test surface of the adaptometer and controlled by the experimenter. The stimulus values, in terms of scale values of the adaptometer were arranged as follows: 0, 10, 50, 100, 200, 300 and up, in steps of 100. Stimulus values were presented in ascending series until one was reached at which light was visible. The irregular low values, 0, 10, and 50, were used for convenience in adjusting the length of the series used for limens which were near the lower end of the scale.

Twenty-five ascending series were used for each limen, which was taken as the average illumination value at which light was reported. Exposure series were varied in length, having from 2 to 8 members; intensities were always increased in regular order except in case of 'doubtful' judgments when a given stimulus value might be repeated. The subliminal starting points of the series were of course unknown to the observers.

The superiority of the second procedure was immediately evident in the reduction in dispersion for the threshold values obtained with its use. The absolute levels of liminal sensitivity were, by the use of the short exposure method, greatly reduced. In some cases the liminal values for the same observer, under the same conditions as used in the first method, were as much as 50 percent lower when measured by the second procedure. This general reduction in thresholds, while it did not eliminate individual differences, served to bring the results of the different observers closer together, since the larger reductions in threshold level occurred for those observers who had shown the highest limens by the first method. The change in threshold was due not to a subjective change in criteria of judgment, but to the brief exposures which did not allow time for adaptation or fatigue in making the measurements.

Limens, and sigmas of the distributions, for monocular and binocular observation of single 2-degree stimulus spots, are given for each observer in Table I. As evident from the individual values, or from the averages, the limens for binocular observation (stimulation of retinal regions $A\alpha$ or $B\beta$) are consistently lower than those for monocular observation (stimulation of single retinal regions, A, α, B, or β). The sigmas are

TABLE I

LIMENS, AND SIGMA OF DISTRIBUTIONS (σ), FOR SINGLE LIGHT STIMULUS OF 2 DEGREES VISUAL ANGLE

(To convert values to millilamberts, multiply by 4.78×10^{-9})

Observer	Retinal Regions Stimulated, and Sigma											
	A	σ	α	σ	$A\alpha$	σ	*B*	σ	β	σ	$B\beta$	σ
B.....	296	99	340	129	240	98	360	103	324	106	272	90
F.....	776	132	684	122	484	96	736	97	814	177	592	93
H.....	620	113	616	118	428	76	1,120	187	904	126	788	76
S.....	436	93	412	147	230	130	416	100	472	141	264	65
T.....	436	132	408	126	276	120	456	136	544	217	318	135
Averages.	512		492		331		617		611		446	

somewhat lower for binocular observation, indicating less irregularity in judgments made with both eyes. As shown in Table II, which gives the actual differences, the standard error of each difference, and the critical ratios for monocular and binocular observation, the difference in favor of greater

TABLE II

ACTUAL DIFFERENCE (*D*), STANDARD ERROR OF THE DIFFERENCE (σ diff), AND CRITICAL RATIOS (D/σ diff) BETWEEN MONOCULAR AND BINOCULAR OBSERVATIONS OF A SINGLE LIGHT SPOT

	Observers				
	B	F	H	S	T
Retinal points:					
A and $A\alpha$					
D.....	56.0	292.0	192.0	206.0	160.0
σ diff.....	27.8	32.6	27.2	31.9	35.6
D/σ diff.....	2.0	8.9	7.0	6.4	4.4
α and $A\alpha$					
D.....	100.0	200.0	188.0	182.0	132.0
σ diff.....	32.4	30.1	24.2	39.2	34.8
D/σ diff.....	3.0	6.6	7.7	4.6	3.7
B and $B\beta$					
D.....	88.0	144.0	332.0	152.0	138.0
σ diff.....	27.4	26.8	40.3	23.0	38.3
D/σ diff.....	3.2	5.3	8.2	6.6	3.6
β and $B\beta$					
D.....	52.0	222.0	116.0	208.0	226.0
σ diff.....	27.8	39.9	29.4	31.0	51.1
D/σ diff.....	1.8	5.5	3.9	6.7	4.4

binocular sensitivity is, in 18 out of 20 cases, significant.[6] In the two cases in which the critical ratios were not large enough to meet the criterion of significance, the difference was in the same direction as that of the significant differences.

The percentage reduction in the stimulus intensity effective for binocular observation, compared with monocular observation, is presented for each observer in Table III. The values

TABLE III

PERCENTAGE REDUCTION IN STIMULUS INTENSITY FOR BINOCULAR OBSERVATION (RETINAL AREAS $A\alpha$ or $B\beta$) OF SINGLE LIGHT SPOTS AS COMPARED WITH MONOCULAR OBSERVATION OF THE SAME SPOTS (RETINAL AREAS A, α, B, β)

Observers	% Reduction for $A\alpha$ (Compared with A & α)	% Reduction for $B\beta$ (Compared with B & β)
B	24	20
F	33	23
H	30	22
S	29	40
T	34	36
Averages	30	28

were calculated by the use of a formula suggested by Laird (**22**).[7] The range in percentage reduction of stimulus values, for the five observers, is from 20 percent to 40 percent, the average being 29 percent.

Table IV presents the data obtained from the use of two

TABLE IV

LIMENS, AND SIGMA OF DISTRIBUTIONS (σ), FOR TWO LIGHT STIMULI OF 2 DEGREES VISUAL ANGLE

Observer	Retinal Regions Stimulated, and Sigma													
	$A\beta$	σ	$B\alpha$	σ	AB	σ	$\alpha\beta$	σ	$A\alpha$-$B\beta$	σ	$AB\beta$	σ	$A\alpha\beta$	σ
B	314	103	340	128	316	89	346	91	232	78	284	96	316	122
F	814	121	784	118	836	171	628	63	464	101	712	114	664	109
H	840	160	880	128	640	101	660	132	388	83	580	105	612	142
S	316	83	302	145	300	93	260	89	312	81	384	115	432	125
T	648	174	572	118	436	144	716	135	500	146	416	100	614	221
Averages	586		575		505		522		379		475		527	

[6] The formula for critical ratios was: Diff./σ diff., where σ diff. $= \sqrt{\sigma^2 \text{ av. } 1 + \sigma^2 \text{ av. } 2}$. The criterion of significance was: Diff./σ diff. > 3.

[7] Laird's formula is: $100 \dfrac{\dfrac{(RL_L + RL_R) - RL_{bin}}{2}}{\dfrac{(RL_L + RL_R)}{2}}$.

RL_L = threshold for left retina; RL_R = threshold for right retina; RL_{bin} = binocular threshold.

2-degree light spots, one on either side of the visual field, 10 degrees from the point of fixation and so arranged that retinal areas having non-identical neurological elaboration could be stimulated. Thresholds are given for the intensity at which both stimuli were simultaneously visible. Limens for retinal areas AB, $\alpha\beta$, $A\beta$, and $B\alpha$, non-corresponding points, are approximately the same as the values for the perception of a single spot seen monocularly (A, α, B or β, in Table I). For both sets of limens, the central projections are such that only one hemisphere is affected by each stimulus. The use of two stimuli, doubling the retinal areas stimulated, is ineffective in lowering the threshold when those areas are not connected with the same central region.

The values in Table IV show that there is no great difference in the limens obtained for retinal areas located on each side of the same eye (AB or $\alpha\beta$) and those for non-corresponding points of the two eyes ($A\beta$ or $B\alpha$). When both stimuli are exposed to both eyes, ($A\alpha B\beta$) the reduced limens demonstrate that binocular observation is as effective in lowering the threshold for the use of two separated stimuli as it was in the case of only one stimulus.

For the limens $AB\beta$, and $A\alpha\beta$, conditions were such that one spot was seen monocularly and the other binocularly. The stimulus seen with both eyes was frequently remarked to be more stable, brighter, and easier to observe than the one seen monocularly. The observers were not aware of the conditions. Both stimuli were often reported at the same intensity, though when one appeared first, it was usually the one seen binocularly. No direct comparison of the two spots was required of the observers, since the stimuli were separated 20 degrees in the visual field and hence would be difficult to judge as to comparative brightness at threshold intensities.

Of interest is the fact that the average limen for the stimulation of three retinal regions, 505 ($AB\beta$ and $A\alpha\beta$), lies between the average limen for two retinal regions, 547 (AB, $\alpha\beta$, $A\beta$, and $B\alpha$), and that for four retinal regions, 379 ($A\alpha B\beta$). The last limen ($A\alpha B\beta$) shows a typical binocular reduction of 30 percent when compared with the average limens for AB,

$\alpha\beta$, $A\beta$, and $B\alpha$. If the reduction in threshold value shown by stimulation of three retinal regions is of significance, it may be an indication of some type of retinal facilitation resulting from simultaneous stimulation of both sides of one retina, or of some interaction between the two hemispheres.

Since the effective intensities for binocular observation are, on the average, 30 percent lower than those for monocular observation, and since the stimulation of non-corresponding points does not lower thresholds, there is evidence of central summation for binocular perception of liminal light stimuli.

The known inverse relationship between stimulus area and threshold intensity under conditions of dark adaptation (summarized by Adams, **3**) made it necessary to determine whether or not stimulus size may influence the reduction in intensity present in binocular observation of a stimulus area. For three observers, a series of monocular (A) and binocular ($A\alpha$) limens was measured for stimuli of 1, 4, and 8 degrees' visual angle, to supplement the data for 2-degree stimuli. The data in Table V, showing the limens and the sigmas of the

TABLE V

MONOCULAR AND BINOCULAR LIMENS, AND SIGMA OF THE DISTRIBUTIONS (σ) FOR STIMULI OF 1, 2, 4, AND 8 DEGREES ANGULAR DIAMETER

Observer	Retinal Region	Angular Diameter							
		1°	σ	2°	σ	4°	σ	8°	σ
B.....	A	1,124	122	296	99	168	42	158	24
	$A\alpha$	788	139	240	98	134	39	99	31
F.....	A	2,668	250	776	132	368	59	274	34
	$A\alpha$	1,476	181	484	96	224	42	129	28
H.....	A	2,060	179	620	113	308	71	214	32
	$A\alpha$	1,452	110	428	76	210	46	177	31
Averages.....	A	1,950		564		281		215	
	$A\alpha$	1,238		384		189		135	

distributions for A and $A\alpha$, for angular diameters of 1, 2, 4, and 8 degrees, indicate that there is a reduction of effective intensity as the size of the stimulus is increased, the largest absolute reduction occurring for the increase from 1 to 2 degrees.[8] For each increase in stimulus size, as shown by the

[8] For 4- and 8-degree stimuli, the low threshold values made it necessary to use a series of stimulus intensities having smaller differences than those for limens of 1- or 2-degree stimuli.

average values in Table V, there is a reduction in liminal values for both monocular and binocular perception. The absolute difference between the average thresholds for monocular and for binocular observation becomes smaller as the stimulus size is increased, but the relative reduction remains approximately the same for each stimulus size. The average values from Table V, when plotted in graphic form, yield curves which are best described by a rectangular hyperbolic function.[9] As the stimulus size is reduced, the intensity necessary for sensation becomes extremely high. The fact that the largest absolute differences between monocular and binocular limens occur for the smaller stimulus areas indicates that with the use of less accurate measuring instruments, the differences would be more easily demonstrated for small stimuli. With the use of extremely large stimuli, aside from the difficulty offered by the varying sensitivity of the peripheral retina, interference of the blind spot, etc., difference between monocular and binocular thresholds would be almost impossible to demonstrate because of the small absolute differences in the limens.

The greater effectiveness of a stimulus having a large angular diameter was demonstrated by the use of two unequal stimuli. A 2-degree stimulus was presented on the left side of the field of vision while a 1-degree stimulus was presented on the right side of the field. When binocular observation was used, the larger of the two stimuli was usually reported first. Two limens were calculated from the data obtained under these conditions: one for the point at which the larger spot was reported and one for the point at which the two were simultaneously visible. Data are given in Table VI. The intensities at which both stimuli were perceived together were lower than the intensities at which, in the previous tests, the 1-degree stimulus, observed binocularly, was reported visible. In other words, a 1-degree stimulus is perceived at a lower intensity when a 2-degree stimulus is presented simultaneously on the

[9] The values obtained (method of least squares) for the function $Y = 1/bX$ showed the relationship between angular diameter of the stimulus (X) and threshold intensity (Y) to be well represented by this formula. The constant (b) for monocular data was .00065; for binocular data, it was .00101.

TABLE VI

LIMENS FOR SIMULTANEOUS PRESENTATION OF UNEQUAL STIMULI (1-DEGREE STIMULUS ON THE RIGHT OF THE VISUAL FIELD; 2-DEGREE STIMULUS ON THE LEFT OF THE VISUAL FIELD)

Observers	Intensity at Which 2-Degree Stimulus was Reported ($B\beta$)	Intensity at Which Both Stimuli were Reported ($A\alpha B\beta$)
B.	356	392
F.	644	1116
H.	744	864

opposite side of the visual field. When conditions were such that the smaller of the two stimuli was seen binocularly while the larger was seen monocularly, the larger 2-degree stimulus remained dominant. Even though the smaller spot (1-degree) was seen with both eyes, it did not equal either in retinal area affected or in apparent brightness the larger spot which was exposed to only one eye. The dominance of the larger stimulus under these conditions might be shown in one of two ways: either the 2-degree spot was invariably reported at a lower intensity than the 1-degree spot, or the two were seen at the same intensity but the larger spot was reported to appear brighter than the smaller.

Since an increase in the area of a light spot reduces the threshold intensity under certain conditions, and since Granit (**16**) and others have demonstrated a considerable amount of interaction between different areas of the same retina, a series of limens was determined for two discrete 2-degree stimuli presented in the upper right quadrant of the visual field. Two such spots should affect the same hemisphere though not necessarily the identical central region; summation, if demonstrated, might be either central or purely retinal. Results for discrete stimuli, separated 3 or 5 degrees from one another, and 10 degrees from the fixation light, are given in Tables VII and VIII. These tables present data for the intensity at which light was first reported and also for the point at which the two stimuli were recognized as discrete. Qualitative reports were analogous to those often given in measuring the two-point limen for cutaneous sensitivity: first a diffuse sensation ('light,' 'cloud,' 'fog') followed by linear extension ('one spot extended,' 'line of light'); then, at a

TABLE VII

LIMENS FOR LIGHT PERCEPTION, AND FOR DISCRIMINATION OF TWO STIMULI AS DISCRETE SPOTS (TWO 2-DEGREE STIMULI, SEPARATED 5 DEGREES)

Observer	Monocular (Retinal Region *A*)		Binocular (Retinal Regions *Aα*)	
	Light	2 Discriminated	Light	2 Discriminated
B	568	784	256	456
F	800	916	552	756
H	614	868	340	576
Averages	660	856	382	596

TABLE VIII

LIMENS FOR LIGHT PERCEPTION, AND FOR DISCRIMINATION OF TWO STIMULI AS DISCRETE SPOTS (TWO 2-DEGREE STIMULI, SEPARATED 3 DEGREES)

Observer	Monocular (Retinal Region A)		Binocular (Retinal Regions *Aα*)	
	Light	2 Discriminated	Light	2 Discriminated
B	452	729	368	580
F	436	1144	480	628
H	484	684	380	492
Averages	457	852	409	566

higher intensity, two points of light connected ('dumb-bell') and finally, a differentiation of two definite light spots. As shown by the data in Tables VII and VIII, binocular observation gives lower limens both for light perception and for discrimination of the two discrete spots than monocular observation, indicating that the summation typical of binocular perception is present. Had there been either retinal or central summation resulting from the use of two stimuli on the same side of the visual field, the limens for light perception should have been lower than those in which light of a single 2-degree stimulus was reported. A comparison of the average values of *A* and *Aα*, from Table I, and the average values for *light* perception in Tables VII and VIII makes it evident that an increase of area by increasing the number of discrete stimulus spots is, under these conditions, not effective in lowering the threshold values.

A number of investigators, including Johannsen and Crook

(**21**) have reported a difference in sensitivity of the two sides of the retina. Usually either the nasal or the temporal regions have been reported more sensitive, but a comparison of the average values in Table I shows that the left halves of each retina (left temporal region A, and right nasal region α) are superior to the right halves of each retina under our conditions. Likewise, the average limen for the stimulation of corresponding points by a stimulus on the right side of the visual field ($A\alpha$) is lower than the average limen for a stimulus on the left side of the visual field ($B\beta$). The differences are not statistically reliable but they indicate, rather than a difference between nasal and temporal regions, a difference between homonymous retinal halves. In terms of cortical projection, excitation of the left hemisphere seems to occur at a lower intensity than the excitation of the right hemisphere. The observers were all right-handed; three of them were right-eyed and two were left-eyed. Franz (**13**) who used the learning of figures as a test, found no consistent superiority of the halves of the retinae which under the commonly accepted theory of left cerebral dominance in right-handed persons should have been superior. On the basis of data from only five observers, there is no justification for concluding that there is any definite superiority of sensitivity for the retinal halves which are in communication with the cerebral hemisphere supposedly dominant, in motor activities, for right-handed individuals.

Conclusions

In order to determine whether or not there is binocular summation in the perception of light by the periphery of the dark-adapted eye, a systematic variation in the conditions of monocular and binocular observation was carried out. Since the peripheral retina reaches a rather stable level of sensitivity after 20 to 30 minutes in the dark, with careful control of the region of the retina used, successive tests made during different periods of observation could be compared. A method consisting of short exposures of discrete stimulus intensities used in ascending series was found satisfactory.

The results of the experimental investigation of thresholds

for light perception by the periphery of the scotopic eye may be summarized briefly:

1. Limens for perception of single light spots of 1, 2, 4, or 8 degrees visual angle, 10 degrees from the point of fixation, are invariably lower for binocular observation than for monocular observation. The stimulus intensity for binocular thresholds is approximately 30 percent lower than that for monocular thresholds. This indicates a definite central summative effect.

2. The use of two light spots, for stimulation of non-corresponding points of the two retinæ, or stimulation of both sides of the same retina, results in liminal intensity values approximating those for monocular thresholds, giving no evidence of summation. When both stimuli are exposed to both eyes, there is an average reduction of 30 percent in the threshold value, indicating again that summation is mediated by binocular observation. When one of two stimulus spots is seen binocularly and the other monocularly, there is a slight reduction in the threshold intensity for simultaneous perception of both spots from that required for perception of both spots by non-corresponding points of the retinæ.

3. Increase in the angular diameter of a stimulus results in decrease of the threshold intensity for the dark-adapted eye. The fact that binocular limens are lower than monocular limens for each stimulus size (1, 2, 4, or 8 degrees) used, indicates that summation is not limited to stimuli of small visual angles. As the stimulus size is increased, however, the difference between monocular and binocular limens, in absolute intensity values, becomes smaller and it is therefore more difficult to demonstrate central summation for large areas.

4. When stimuli of unequal area are presented simultaneously on either side of the visual field, for binocular observation, the larger area is dominant: either it is seen at a lower intensity than the smaller stimulus, or it is reported to appear brighter and more stable at the threshold intensity than the smaller stimulus. The reduction in intensity resulting from central summation for binocular observation of a 1-degree spot is not sufficient to overcome the dominance of a larger 2-degree spot presented monocularly at the same time.

5. Increase in area by the use of two discrete stimulus spots, separated 3 or 5 degrees, presented in the same quadrant of the visual field, gives no reduction in threshold values for monocular perception. A summative effect for both light perception and for discrimination of discrete stimuli as separate spots of light was demonstrated for binocular observation of discrete stimuli.

6. Although there was some evidence of a superiority of left homonymous retinal halves, it was not sufficient to indicate decisively that there is left cerebral dominance in the visual perception of right-handed individuals

Theoretical Discussion

A consistent reduction of about 30 percent in the intensity for binocular thresholds as contrasted with monocular thresholds, demonstrated only when corresponding points of the retinæ, and hence identical central regions, were stimulated has been interpreted as an indication of central summation.

Lower binocular thresholds have been attributed by Müller (**26**) to 'greater ease' of binocular observation, but under conditions in which binocular fixation is used, this would not suffice to explain a significant reduction in the binocular threshold: binocular summation may in itself contribute to the greater ease of observation with both eyes.

Although under normal conditions there is usually superiority of binocular acuity, as shown by Ferree, Rand and Buckley (**12**), the measurements made under the conditions of the present study, with advanced dark adaptation and peripheral observation probably involve photochemical sensitivity rather than resolving power of the eye.[10]

A number of investigators, including Hecht (**17**), Heinz and Lippay (**18**) and Lasareff (**23**) have emphasized the importance of the number of excited retinal elements in various visual functions. Hecht uses this as a basis for an explanation of acuity as a function of the illumination, differences in intensity being assumed to be correlated with

[10] Reeves (**30**, p. 142) states that the threshold for dark adaptation gives the absolute photochemical reactivity of the retina.

differences in the number of activated retinal elements. Lohmann, (**24**) who suggested that summation in binocular observation is a matter of increased retinal excitation, has attempted to apply the inverse relationship between stimulus area and threshold intensity in an explanation of lower binocular thresholds. According to Lohmann, binocular stimulation results in the excitation of non-identical retinal regions which is equivalent to an increase in stimulus area and thus accounts for the lower thresholds. This hypothesis would imply that, centrally, it is the stimulation of non-identical rather than identical regions which mediates lower thresholds: therefore, excitation of non-corresponding points should be as effective as stimulation of corresponding points of equal area. It has been shown, however, that when equivalent retinal areas are stimulated, summation does not occur unless some of the retinal elements stimulated are corresponding points exciting identical, or nearly identical, central regions.

The concept of number of excited sensory elements, inadequate to explain central summation, would be equally inadequate to explain the lowered threshold values resulting from increase in stimulus area. If merely the number of excited elements were significant, it would be immaterial whether the increase in area occurred by enlarging a single area or by addition of another discrete area. The decrease in threshold intensity accompanying the increase in stimulus area, demonstrated by Piper (**29**) and others has been found also by Adrian and Matthews (**4**, Part 1) in studies of action currents in the excised eyes of eels. An increase in area is equivalent to raising the intensity of the stimulus. Impulses may be summated on the way to the optic nerve, an effect described by Granit (**16**) as synaptic reaction, or 'spatial summation' effected through horizontal cells, amacrines, and cells of the ganglionic layer which furnish lateral junctions. Granit has defined the summative effects such as those of Adrian and Matthews as 'The physiologic expression of the anatomic convergence of a number of paths to a common neuron.' It seems possible that the summation

resulting from binocular stimulation is likewise to be explained on the basis of the union of impulses from several regions. It is certain that the summative effects demonstrated for binocular stimulation cannot be accounted for without reference to the central regions involved.

Since it becomes more difficult to measure differences between monocular and binocular thresholds as the stimulus area is increased, on account of the small absolute difference between the two thresholds, it is possible that maximum stimulation from one eye would prevent a reduction in threshold intensity by additional excitation of the same central region by impulses from the other eye. This may explain why summative effects have not been satisfactorily demonstrated except for minimal intensities, and have frequently been reported as not existing under conditions of photopic vision.

The manner in which summation occurs remains a matter of speculation until more complete information concerning the combination of impulses within the central nervous system can be secured. If there is a spatial representation of peripheral retinal regions in the visual cortex, or in sub-cortical areas, stimulation of corresponding retinal points, resulting in concentration of central excitation effected through divergent peripheral regions, may account for summation.

(Manuscript received May 31, 1934)

References

1. Abney, W. deW., & Watson, W., The threshold of vision for different colored lights, *Phil. Trans. Roy. Soc.* 1916, **216**, 91–128.
2. Achmatov, A. S., Eine experimentelle Untersuchung der Dunkeladaptationsgleichungen, *Arch. f. d. ges. Physiol.* 1926, **215**, 10–18.
3. Adams, D., Dark adaptation (a review of the literature), Medical Research Council, special report series, No. 127. London: His Majesty's Stationery Office, 1929, Pp. 138.
4. Adrian, E. D., & Matthews, R., The action of light on the eye. *J. Physiol.*, 1927, **63**, 378–411 (Part I); 1927, **64**, 279–301 (Part 2); 1928, **65**, 273–298 (Part 3).
5. Cook, T. W., Binocular and monocular relations in foveal dark adaptation, *Psychol. Mono.*, 1934, **45**, No. 3, Pp. 83.

6. Behr, C. v., Der Reflexcharakter der Adaptationsvorgänge, insbesondere der Darkadaptationen und deren Beziehungen zur topischen Diagnose und zur Hemeralopie, *Arch. f. Ophth.*, 1910, **75**, 201–282.
7. Best, F., Untersuchung über die Dunkelanpassung des Auges mit Leuchtfarben. *Zsch. f. Biolog.*, 1917, **68**, 111–145.
8. Blanchard, J., The brightness sensibility of the retina. *Phys. Rev.*, 1918, Ser. 2, **11**, 81–99.
9. DeSilva, H. R., & Bartley, S. H., Summation and subtraction of brightness in binocular perception, *Brit. J. Psychol.*, 1930, **20**, 241–250.
10. Downey, J. W., Determination of minimum light sense and retinal dark adaptation, *Amer. J. Ophth.*, 1919, **2**, 13–20.
11. Feilchenfeld, H., & Loeser, L., Ueber die Beeinflussung einer Lichtempfindung durch eine andere gleichzeitige Lichtempfindung, *Arch. f. Ophth.*, 1905, **60**, 97–115.
12. Ferree, C. E., Rand, G., & Buckley, Study of ocular functions with special reference to the look-out and signal service of the navy, J. Exper. Psychol., 1920, **3**, 347–356.
13. Franz, S. I., Inadequacy of the concept of unilateral cerebral dominance in learning, J. Exper. Psychol., 1933, **16**, 873–875.
14. Fröbes, S. J., Lehrbuch der Experimentellen Psychologie, Freiburg, 1923.
15. Graham, C. H., An investigation of binocular summation, *J. Gen Psychol.*, 1930 **3**, 494–510 (Part 1); 1931, 5, 311–328 (Part 2).
16. Granit, R., The retina as a nervous center, *Arch. Ophth.*, 1931, **6**, 104–111.
17. Hecht, S., The nature of the photoreceptor processes, Foundations of Experimental Psychology. Worcester: Clark University Press, 1929, 216–269.
18. Heinz, M., & Lippay, F., Ueber die Beziehungen zwischen der Unterschiedsempfindlichkeit und der Zahl der erregten Sinneselemente, *Arch. f. d. ges. Physiol.*, 1927, **218**, 437–447.
19. Helmholtz, H. v., A Treatise on Physiological Optics (Ed. by J. P. C. Southall). Ithaca, N. Y.: Optical Soc. Amer., Vol. 2, pp. viii–480.
20. Helson, H., & Fehrer, E. V., The role of form in perception, *Amer. J. Psychol.*, 1932, **44**, 79–102.
21. Johannsen, D. E., & Crook, M. N., Differential adaptation of the two sides of the retina, *J. Gen. Psychol.*, 1930, **3**, 307–312.
22. Laird, D. A., Studies relating to the problem of binocular summation, J. Exper. Psychol., 1924, **7**, 276–290.
23. Lasareff, P., Die mathematischen Theorien der Dunkeladaptation des menschlichen Auges, *J. f. Psychol. u. Neur.*, 1926, **32**, 232–252.
24. Lohmann, W., Kritische Studien zur Lehre von der Adaptation, *Arch. f. Augenhk.*, 1918, **83**, 275–292.
25. Messmer, —, Ueber die Dunkeladaptation bei Hemeralopie, *Zsch. f. Sinnesphysiol.* 1907, **42**, 83–100.
26. Muller, E., Die monokulare und binokulare Reizschwelle der dunkeladaptierten Augen, *Arch. f. d. ges. Physiol.*, 1921, **193**, 29–38.
27. Peter, L. C., The Principles and Practise of Perimetry. Philadelphia: Lea & Febiger, 1923. Pp. xii–281.
28. Piper, H., Ueber Dunkeladaptation, *Ztsch. f. Psychol.*, 1903, **31**, 161–214.
29. ——, Ueber die Abhängigkeit des Reizwertes leuchtender Objekte von ihrer Flächen, *Zsch. f. Psychol.*, 1903, **32**, 98–112.

30. Reeves, P., The effect of size of stimulus and exposure time on retinal threshold, *Astrophys. J.*, 1918, **47**, 141–145.
31. Roelofs, C. O., & Zeemann, W. P. C., Zur Frage der Binokularen Helligkeit und und der binokularen Schwellenwerte, *Arch. f. Ophth.*, 1914, **88**, 1–27.
32. Sherrington, C. S., On binocular flicker and the correlation of activity of 'corresponding' retinal points, *Brit. J. Psychol.*, 1904, **1**, 26–60.
33. Stargardt, K., Ueber Störungen der Dunkeladaptation, *Arch. f. Ophth.*, 1910, **73**, 77–162.
34. Wölfflin, E., Ueber Dunkeladaptation von fovealen und parafovealen Netzhautpartien, *Arch. f. Ophth.*, 1910, **76**, 464–477.

THE RELATIONSHIP BETWEEN TIME SPENT IN THE CULS-DE-SAC OF A STYLUS MAZE AND SPEED OF ELIMINATION

BY HENRY N. PETERS AND JOHN A. McGEOCH

University of Missouri

Problem

The operation of the time factor in the acquisition of human skill has previously been studied only with reference to the performance of the total act. In the case of the maze, for instance, the only *time* that has attracted any notice has been the interval between the starting point and the entrance into the goal box. It was the purpose of this experiment to make an analytical study of the time factor in the learning of a skilled act by investigating the time relations which obtain *within* the learning process.[1] The experimental method was particularly devised to make possible a determination of the relation between varying intervals of time spent in the blind alleys of a stylus maze and the difficulty of elimination.

The procedure usually employed in studies of the influence of a specific factor upon learning is that of dividing the subjects into groups and forcing upon each group of learners a variation of the factor to be studied. Treatment of the data is then a matter of comparing the results obtained for different groups. In the present experiment the subjects were neither divided into groups nor were varying conditions forced upon them. Careful records were taken for each subject and the condition to be investigated, *i.e.*, the interval of time spent in the alleys, obtained in this fashion was then studied for its own sake.

Method.—The maze pattern used in this experiment was an exact reproduction of Warden's 10-alley maze (Pattern 1).[2] The alleys were cut ¼ inch in width out of bake-

[1] For a previous report of one phase of the findings of this study, see J. A. McGeoch and H. N. Peters, An all-or-none characteristic in the elimination of errors during the learning of a stylus maze, J. Exper. Psychol., 1933, **16**, 504–523.

[2] C. J. Warden, Primacy and recency as factors in cul-de-sac elimination in a stylus maze, J. Exper. Psychol., 1924, **7**, 98–116.

lite and mounted on a base of smooth bakelite. Brass contact points at the appropriate places and with the necessary electrical connections made it possible to record exactly the length of time spent by the subject in each alley. For a detailed description of the apparatus used in the experiment, the reader is referred to the previous report by the same authors in this journal.[3]

Thirty college students of an average age of twenty years, none of whom had seen a stylus maze before, were used as subjects. The subject was seated behind a screen which totally concealed the maze and recording apparatus, and the customary maze instructions were read to him. The learning process was continued until the criterion of three perfect trials in succession was reached. No attempt was made to hinder the subject from retracing in the true path of the maze; however, for this study retracings were not counted as errors. As the subject ran the maze, the experimenter recorded the number of each blind alley, as it was entered, directly above the corresponding indenture made by the recording pen on the moving band of paper. This record made possible the determination, to a fifth of a second, of the time spent by each subject in each cul-de-sac entrance. The beginning and end, as well as the total time in seconds, of each trial were also designated on these records.

The organization of the raw data consisted of the task of counting and recording in tabular form the intervals of time spent in each cul-de-sac entrance and arranging them for each subject according to the trial and the serial number of the blind alley. In addition, these tabular sheets contained the usual maze data. From these sheets, of which there was one for each subject, master sheets were compiled to indicate the average entrance time for each subject for all alleys and for each of the ten individual alleys, with corresponding figures for various criteria of difficulty of elimination, such as, number of entrances, number of trials, and total time. Similar master sheets were made for each of the ten individual blind alleys, showing the average entrance time for each alley, with corresponding figures for the various criteria of its difficulty of elimination.

Results

I. *The Relation between the Length of Time Spent in the Blind Alley and Difficulty of Elimination.*—Coefficients of correlation[4] constitute the chief evidence presented in this paper for a functional relationship between time per error and difficulty of elimination. The two variables represented in each correlation are mean blind alley time and a measure of rate of learning, or of difficulty of elimination. These two series of variables were determined for all the alleys taken together and for each blind alley individually. In the former case, the blind alley time was found for each subject by dividing the total time spent in all blind alleys by the total number of entrances; this was correlated with measures of rate of

[3] J. A. McGeoch and H. N. Peters, *op. cit.*

[4] All the coefficients of correlation listed in the following results were obtained by the Spearman rank-difference method and transposed into Pearson *r*'s. In transposing, the size of the coefficients was increased very little, at the most three one-hundredths.

learning the total maze pattern. Table I shows the resulting coefficients for three measures of learning. The number of cases here is the number of subjects (30). The coefficients presented in this table (and another set to be presented later in Table III) are considered more important than those for the individual alleys, because the extraneous factors operating in the case of any one alley should balance out when all ten are considered in composite.

These figures show an inverse relation between the length

TABLE I

COEFFICIENTS OF CORRELATION BETWEEN MEAN TIME PER SUBJECT IN ALL BLIND ALLEYS AND THREE MEASURES OF LEARNING OF THE TOTAL PATTERN

Measures of Learning		
Number of Trials	Number of Errors	Total Time
$r = -0.59 \pm .08$	$r = -0.45 \pm .10^{5}$	$r = +0.20 \pm .12$

of time spent in blind alleys and the difficulty of learning the maze. In other words, the longer the mean cul-de-sac time, the greater the ease of learning the maze pattern. Since the first two of the correlations are fairly high numbers, and are more than four times their P.E., the tendency which they show may be accepted as valid. Investigation of the variables of the third coefficient will convince the reader that its implications are in accord with the others. In the first place, there is a spuriousness in this correlation. Both of the variables have to do with the relative length of time spent in the learning of the maze. If a subject spends long periods of time in the blind alleys, these periods accordingly lengthen his total time. In the second place, and as will be brought out later in this paper, there is practically a constant ratio between time in alleys and time in the true path for all subjects, *i.e.*, a subject spending long periods of time in the alleys also spends relatively long periods of time in the true path. Taking these two factors into consideration, the remarkable thing is that this coefficient is not a high positive one.

[5] This is not a case of spurious index correlation, because the number-of-errors factor, which is denominator in the average time per entrance variable, is a variable quantity.

Table II shows correlations computed for each of the ten alleys instead of all of the alleys taken together. The first column shows the number of the alley; the second column, the coefficients of correlation between mean time per entrance and the number of entrances into the alley during learning; the third column, the coefficients between mean time per

TABLE II

COEFFICIENTS OF CORRELATION BETWEEN MEAN TIME PER ENTRANCE AND THREE CRITERIA OF DIFFICULTY FOR INDIVIDUAL ALLEYS

Alley No.	*r* between Mean Time per Entrance and		
	Number of Entrances[6]	Trials to Eliminate	Trials to Five Perfect Runs
I.....	−.35±.11	−.41±.10	−.05±.12
II.....	+.09±.12	+.17±.12	+.15±.12
III.....	−.33±.11	−.41±.10	−.21±.12
IV.....	−.59±.08	−.41±.10	−.54±.09
V.....	−.06±.12	−.16±.12	−.12±.12
VI.....	−.42±.10	−.39±.10	−.19±.12
VII.....	−.19±.12	−.36±.11	+.07±.12
VIII.....	+.31±.11	+.21±.12	+.33±.11
IX.....	−.51±.09	−.39±.10	−.38±.10
X.....	+.01±.12	−.25±.12	−.15±.12

[6] See footnote 5.

entrance and the number of trials (for the total maze pattern) to eliminate the alley, *i.e.*, number of trials beyond which the alley was not entered; the fourth column, the coefficients between mean time per entrance and the number of trials (for the total maze pattern) to eliminate the alley for five consecutive runs. This last criterion for an individual alley of avoidance for five successive trials was deemed important, because, as the records show, an alley was often eliminated for five or more successive trials early in learning only to be subsequently entered for several trials before being eliminated a second time.

These figures, on the whole, support those for all the alleys treated as one. The relation, if any, between blind alley elimination and entrance time would be expected to show itself less in the case of the separate alleys because of the extraneous factors, especially serial position effects, operating differently for the different alleys. Alley II and alley VIII

are the only ones showing consistent positive correlations; those for VIII are the only large positive correlations, and only one of these is three times its P.E. There is no apparent reason for the diversion of coefficients for this alley from the general tendency. Alley VIII was one of the most difficult in the whole maze for the subjects to eliminate.

The *r*'s between mean time per entrance and trials to five perfect runs tend to be of less significance than the others. This is perhaps due to the unreliability of five perfect trials as a measure of relative difficulty. As the data show, a subject quite frequently succeeded in avoiding an alley for five or six trials only to return to it with even greater consistency later on. However, the variations in the coefficients of Table II are not a result of chance entirely. The remarkable feature of the table is the relative consistency of the figures opposite the individual culs-de-sac. When the coefficients in each column of Table II were correlated with those in the other two columns, the following *r*'s were obtained: correlating column two with column three, $r = +0.86 \pm .06$; correlating column two with column four, $r = +0.84 \pm .06$; correlating column three with column four, $r = +0.74 \pm .12$. These correlations are too high and consistent to be caused by chance factors. The conclusion is that the height of the correlations between mean time per entrance and the criteria of difficulty employed vary consistently with the particular alleys.

In all the coefficients given thus far, showing the relationship between the two variables, mean time per entrance and difficulty of alley elimination, the individual subjects were the separate cases. Coefficients have also been calculated between the same variables in terms of means of the thirty subjects' scores, the alleys constituting the separate cases. In the first case, $N = 30$; in the second, $N = 10$. The coefficients presented in Table III were obtained by the latter procedure.

TABLE III

COEFFICIENTS OF CORRELATION BETWEEN MEAN TIME PER ENTRANCE PER ALLEY AND THREE MEASURES OF DIFFICULTY OF ELIMINATION

Measures of Difficulty		
Number of Trials	Number of Errors	Total Time
$r = -0.63 \pm .13$	$r = -0.56 \pm .15$	$r = -0.43 \pm .17$

The mean time per entrance used in this table was obtained by dividing the total amount of time spent by all thirty subjects in the alley by the total number of entrances made by all thirty subjects into the alley. The number of trials is the total of the maze trials each subject took in finally eliminating the blind alley; number of errors is the total of all subjects' entrances into the blind alley; total time is the sum of the thirty subjects' time in the alley.

These coefficients are larger than the corresponding ones given above (see Table I). And it will be observed that in this instance the correlation with total time turns out decidedly negative. In the former instance (Table I), mean error-time was correlated with the *total* time of learning; here mean error-time is correlated *only* with blind alley time in each case. Thus, here the true-path time is not involved; however, the same spuriousness pointed out in the preceding coefficient exists here. It will be remembered that this spuriousness influences the correlation in the positive direction. The significant thing is that, in spite of this fact, the coefficient is still negative.[7]

The measure of the time spent in each blind alley used in obtaining the above results was the mean of all entrances into each single alley from the beginning to the end of learning. In the case of almost every alley (for each subject) a great many entrances were involved. Curves drawn for the time of these entrances show a tendency to drop from a high, initial level to a constant level which is held to the end of learning.[8] In light of these curves, the following question arose: Which entrances are more influential in determining

[7] The correlation coefficients presented in this paper represent a selection of the most significant from a much larger number that were calculated. For example, the variable mean time per entrance was correlated wherever possible with the four measures of difficulty, number of trials, number of trials to five perfect runs, total time, and total errors, although the resulting coefficients do not appear in the included tables. The greatest number of omitted correlation coefficients, however, are those calculated between mean and absolute time of first entrance into a blind alley and the four measures of rate of elimination. These coefficients all turned out very near zero.

[8] J. A. McGeoch and H. N. Peters, *op. cit.* In this paper three methods were used in plotting the curves for time spent in blind alleys. Here, the reader is referred particularly to the curves drawn according to Method II.

the negative relationship found here, the early or the later ones? According to the curves drawn, the initial, orientation period is over and the constant level reached at approximately the mid-point. As an answer to the question, the mean entrance times for the first half and for the second half of the entrances were determined, and the same correlations which were computed with the mean of all entrances as a variable were computed again separately with the means for the first and second half of the elimination series substituted for the mean of all.

Table IV presents the correlations found between the two

TABLE IV

COEFFICIENTS OF CORRELATION BETWEEN TWO MEAN TIMES, FIRST AND SECOND HALVES, PER SUBJECT IN ALL BLIND ALLEYS AND THREE MEASURES OF LEARNING OF THE TOTAL PATTERN

Mean Time	Measures of Learning		
	Number of Trials	Number of Errors	Total Time
First Half.....	$r = -0.50 \pm .09$	$r = -0.39 \pm .10$	$r = +0.23 \pm .12$
Second Half.....	$r = -0.51 \pm .09$	$r = -0.29 \pm .11$	$r = +0.32 \pm .11$

mean entrance times and measures of learning the total maze pattern. For these coefficients, $N = 30$. Table IV corresponds to Table I, above, for the mean of all entrances. The implication of these coefficients, except for the one involving number of trials, is that the first half of the entrances contributes more than the second to the inverse relation.

Table V shows the correlations found when the alleys

TABLE V

COEFFICIENTS OF CORRELATION BETWEEN TWO MEAN TIMES, FIRST AND SECOND HALVES, PER ENTRANCE PER ALLEY AND THREE MEASURES OF DIFFICULTY OF ELIMINATION

Mean Time	Measures of Difficulty		
	Number of Trials	Number of Errors	Total Time
First Half.....	$r = -0.73 \pm .10$	$r = -0.79 \pm .08$	$r = -0.67 \pm .12$
Second Half.....	$r = -0.45 \pm .17$	$r = -0.22 \pm .20$	$r = -0.02 \pm .21$

are taken as the separate cases (instead of the subjects). In this table, $N = 10$. Table V corresponds to Table III, above, for the mean of all entrances.

As reference to Tables I and III will reveal, the coefficients found when the two halves of the total blind alley entrances are treated separately support the coefficients for the two halves treated together. The coefficients for the first half of the entrances show a stronger negative tendency than do those for the second half; and the coefficients given in Table V (where the variables are considered in terms of means of the thirty subjects' scores, the alleys constituting the separate cases) for the first half of the entrances are larger negative figures than any found for the means of all entrances.

Correlations were run between the means of first and second halves of total blind alley entrances and measures of difficulty for the individual alleys; these measures were the number of entrances, number of trials to eliminate completely, and number of trials to five perfect runs; and the resulting coefficients correspond to those given for the means of all entrance times in Table II. The coefficients for both halves agree closely with those shown in Table II. Those for the first half show slightly larger negative correlations than those for the second half.

In answer to the question, which half of the blind alley entrances is more influential in producing the negative correlation between time and difficulty, we may conclude that, although both first and second halves contribute, the influence of the first is greater than that of the second.

Although the above results are sufficient to warrant the conclusion of an inverse relationship between time spent in the alleys and difficulty of elimination, the fact of this inverse relation can be shown in ways more detailed than that of correlation coefficients. Table VI shows groupings of the subjects into steps, according to mean time spent in alleys per entrance, with the corresponding mean numbers of errors and mean numbers of trials to three perfect runs. Column one includes the number of subjects falling into each group; column two, the mid-point of each half-second interval con-

TABLE VI

GROUPINGS OF SUBJECTS INTO STEP INTERVALS OF ½ SECOND, ACCORDING TO MEAN TIMES SPENT IN ALLEYS PER ENTRANCE

No. of Cases	Mean Time per Entrance	Mean Number of Errors	Mean Number of Trials
3	2.5	276	82
7	3.0	266	80
4	3.5	179	46
6	4.0	161	49
2	4.5	138	39
3	5.0	176	52
3	6.0	165	40
1	8.0	231	51
1	11.0	99	24

cerned; column three, the mean of the total errors of the subjects in the group; column four, the mean number of trials.

This table roughly shows the tendency of errors and trials to decrease as mean time per entrance increases.

II. *The Relation between Total Time Spent in the True Path and Total Time Spent in Blind Alleys.*—In determining the proportions of the total learning time spent in blind alleys and in the true path, a rather surprising fact was discovered, a fact which could be ascertained only by a method similar to that used in this study.

Investigation of this relationship between true path time and blind alley time is important on two counts. In the first place, the question of whether or not the subjects differ with reference to proportion of total time spent in blind alleys becomes an urgent one in light of the above correlation coefficients indicating a difference in the efficacy of brief and lengthy intervals of time in blind alleys. The subjects do differ in mean entrance times. Do they differ in proportion of total learning time spent in blind alleys? In the second place, this relationship is important as a quantitative comparison of that portion of the learning activity taken up with right responses with that portion of the learning activity taken up with wrong responses. Such a direct comparison has never been made; and the method used in this study lends itself easily to the problem.

Table VII shows the total time taken by each subject to

TABLE VII

PERCENTAGES OF THE TOTAL TIME OF LEARNING SPENT IN THE TRUE PATH

Subject	Total Time in Seconds	% of Total Time in True Path
1	1612	.666
2	690	.498
3	3173	.646
4	1230	.695
5	807	.611
6	1158	.604
7	951	.646
8	1699	.600
9	2040	.560
10	1110	.734
11	1888	.671
12	1309	.674
13	1554	.645
14	1565	.684
15	1766	.577
16	3800	.512
17	1272	.586
18	2074	.651
19	2755	.626
20	4143	.638
21	1997	.588
22	3115	.698
23	1270	.562
24	1151	.644
25	2381	.591
26	2287	.579
27	2553	.655
28	2609	.577
29	4172	.550
30	3677	.657
Mean %		.620±.054

learn the maze to a criterion of three perfect runs and the percentage of this time spent in the true path.

Since the subjects who spent relatively long intervals of time per entrance into the blind alleys learned the maze with greater ease, it might be expected that these same subjects would have spent relatively greater proportions of their total time in the blind alleys than did the subjects who spent short intervals of time in blind alleys. Although, as the table reveals, the total times vary extensively, the percentages of the total time spent in the true path (and, therefore, in the blind alleys) approximate a constant for all subjects. The implication of this discovery is that the relative time of errors operates independently of the sheer frequency of the errors, an implication also shown by the correlation coefficients presented earlier in the paper. These thirty subjects varied widely in

the number of trials and errors made during learning of the maze, but the proportion of time they spent making errors was a constant. Whether or not this constant proportion of error-time to total time holds for other kinds of learning, or even other mazes, is unknown.

III. *The Influence of Serial Position of Blind Alley.*—The question of serial position effects has already been touched upon. It will be remembered that a fairly high positive correlation was found, in the case of the ten alleys, between the coefficients of correlation expressing the relationship between time interval and three criteria of difficulty of elimination. This result points toward an influence of serial position upon the rate of the elimination of an alley. In view of the well known results on serial position effects in memory materials, the problem bears directly on the comparative nature of memory and skill. Warden, using the same maze employed here, reports that primacy and finality operate in maze learning as they do in the learning of nonsense syllables.[9] As far as Warden goes in presenting his data in the paper, the figures obtained in this study agree with his. The figures for serial position, however, warrant another conclusion than his when they are viewed in more detail.

Table VIII gives the serial position effects for the ten blind alleys. In the first column are the numbers of the blind alleys; in the second column, the total number of entrances into the

TABLE VIII

SERIAL POSITION EFFECTS FOR THE TEN ALLEYS WITH RESPECT TO NUMBER OF ENTRANCES, TOTAL TIME, AND MEAN TIME PER ENTRANCE

Alley No.	Number of Entrances	Total Time	Mean Time per Entrance (in sec.)
I	414	2,019	4.88
II	209	1,035	4.95
III	666	2,378	3.57
IV	1,038	4,234	4.08
V	340	1,362	4.01
VI	833	2,944	3.53
VII	881	3,578	4.06
VIII	275	1,347	4.90
IX	978	3,378	3.45
X	369	1,344	3.64

[9] C. J. Warden, *op. cit.*

respective blind alleys; in the third column, the total time, in seconds, spent in each blind alley; in the fourth column, the mean time per entrance spent in each of the blind alleys. The figures composing this table were obtained from the group of thirty subjects' records taken as a whole. The criterion of learning the maze was three perfect runs in succession. The figures in the table include all entrances into the culs-de-sac, from beginning to end of learning.

In comparing these results with Warden's we are concerned with only the first two columns of Table VIII. Means, computed as Warden did, for the number of entrances into the first three blind alleys, the middle four blind alleys, and the last three blind alleys support his conclusion. These figures are 429.6, 773.0, and 540.6. But by this method of presentation the averages mask an irregularity obtaining among the separate blind alleys, and an irregularity as important as, if not more important than, the averages themselves. According to the figures in column two of Table VIII, alleys I, II, V, VIII, and X are eliminated with relative ease; alleys IV, VI, VII, and IX, with relative difficulty. If the curve is drawn, no regular primacy-finality effect is evident.

The figures in column three of Table VIII, for total time, roughly parallel those of column two. For total time, as for total errors, the striking thing is the irregularity. If these figures are grouped into averages, however, a primacy-finality effect can also be demonstrated for time. The averages for the first three, the middle four, and the final three blind alleys are: 1,810.6, 3,029.5, and 2,023.0. With total time, as with total errors, the three groups of alleys arrange themselves in order of increasing difficulty as follows: the first three, the last three, and the middle four. Thus, for this maze, the primacy effect is stronger than the finality.

Regarding only the first three columns of this table, one may conclude that the serial position effects found in maze learning are different from those found in memorizing of verbal lists. Although there is a general primacy-finality effect as shown by the average values for the first, last, and middle blind alleys, the effect is not a regular one.

There is probably no one serial position curve for mazes, but rather a different one for different mazes. The nature of the maze pattern itself influences the difficulty of the individual blind alley. On purely *a priori* grounds one would not expect to find a serial position curve for culs-de-sac corresponding to the serial position curve for nonsense syllables. The culs-de-sac are parts of a unified whole, the maze pattern; there is no such objective whole, or pattern, in a list of nonsense syllables.

The fourth column of Table VIII was included by way of a confirmation of the principal finding of the experiment reported in this paper: an inverse relationship between entrance time and difficulty of blind alley elimination. The tendency is for mean time per entrance to be high, relatively, when the figures for number of entrances and total time are low, and for mean time per entrance to be low when the other two are high. The curve for mean time per entrance (when drawn from the figures in the table) looks like the curve for number of entrances inverted. It is also interesting, as further evidence still, to note that the average values for the first three, the middle four, and the last three mean times per entrance are: 4.47, 3.92, 4.00. When arranged in order of increasing size, these average values fall just opposite to the above mentioned order of increasing number of errors and trials.

Discussion

We have found that in learning a stylus maze the general tendency is for subjects who spend longer intervals of time in culs-de-sac to eliminate them with greater ease than do subjects who spend shorter intervals of time in the culs-de-sac. This result has been checked by several different methods.

The problem of time interval in blind alleys may be viewed as one of the relationship between time interval and the rapidity of formation of successive associations. In learning a maze, the subject associates, or connects, the various sections of the true path; the times spent in culs-de-sac constitute the intervals between the items to be associated. Other factors may also be operating, such, for example, as remote associa-

tion, whereby the organization of the total pattern is made more compact, but within the limits of our present data we may treat the problem as one of direct successive association.

The general problem of the influence of the length of the time intervals between items upon their successive association has been investigated in both human and animal psychology. With memory materials and human beings as subjects, Bergström [10] found that in successive association the lengthening of the time interval was decidedly beneficial, while Wohlgemuth [11] and Froeberg [12] found that it was detrimental to the association of the two items. Wolfle's study of the influence of time interval in establishing a conditioned withdrawal of the finger found that "with an interval series in steps of .25 seconds, beginning with the unconditioned stimulus (shock) .50 seconds *before* the conditioned stimulus (sound), and ending with the shock 1.50 seconds after the sound, the mean percentage of conditioning for the separate intervals was 10, 13, 19, 29, 37, 22, 14, 12, and 8." [13] The optimal interval, at which 37 percent of conditioning was reached, was .50 seconds with the sound stimulus first. With rats in the alternation problem-box, Yarbrough [14] found that "upon the basis of the number of trials necessary to make the association, the difficulty of the problem steadily increases with the increase of the time interval." [15] The most recent

[10] J. A. Bergström, Effect of changes in the time variables in memorizing, together with some discussion of the technique of memory experimentation, *Amer. J. Psychol.*, 1907, **18**, 206–238.

[11] A. Wohlgemuth, Simultaneous and successive association, *Brit. J. Psychol.*, 1915, **7**, 434–452.

[12] S. Froeberg, Simultaneous versus successive association, *Phychol. Rev.*, 1918, **25**, 156–163.

[13] H. M. Wolfle, Time factors in conditioning finger-withdrawal, *J. Gener. Psychol.*, 1930, **4**, 372–378.

[14] J. U. Yarbrough, The influence of time interval upon the rate of learning in the white rat, *Psychol. Monog.*, 1921, **30**, No. 135, pp. 1–52.

[15] Results similar to these have been obtained in other experiments on the problem of time interval in association. See, e.g., J. A. Bigham, Studies from the Harvard Psychological Laboratory: A. Memory, *Psychol. Rev.*, 1894, **1**, 34–38; R. M. Ogden, Untersuchungen über den Einfluss der Geschwindigkeit des lauten Lesens auf das Lernen und Behalten von sinnlosen und sinnvollen Stoffen, *Arch. f. d. ges. Psychol.*, 1903, **2**, 93–189; H. A. Carr, Length of time interval in successive association, *Psychol. Rev.*, 1919, **26**, 335–353; H. A. Carr and A. S. Freeman, Time relationships in the forma-

investigation of time interval in successive association is that of Guthrie.[16] With nonsense syllables and nonsense figures as materials and the intervals 4.93, 3.33, 2.55 seconds he found an 'increase in associative strength as the interval increased.' With the exception of Wolfle's experiment, which shows a beneficial influence upon conditioning of lengthening the interval separating the two stimuli from zero to .50 seconds, the studies of Bergström and of Guthrie are the only ones in fundamental agreement with the results of this experiment. And in these two studies memory material was used in which the items were disparate, unrelated, ones. In comparing studies by widely different methods it is difficult to give agreement or disagreement a systematic interpretation.

The results reported in this paper may be interpreted in terms of the law of effect. May not the time interval in the blind alleys bear somewhat the same relation to effect as Carr has maintained for frequency?[17] That is, just as frequency, or repetition, allows the fixating or eliminating influence of effect to operate, so also does the time interval in the case of blind alleys in the maze. The finding that alleys are eliminated in inverse relation to the length of time spent in them indicates (as other recent experiments have) that the doctrine of frequency is in need of modification. According to the results presented here, the lower frequencies (and longer entrances) tend to be correlated with faster elimination; and it appears that it is the time spent, as well as frequency, which matters. A factor of *temporal magnitude* or *temporal extent* must be added to those influencing learning.[18] It is probable that this time factor is not a directly acting one, but that a given temporal extent favors verbalization, orientation, exploration with attendant completeness of information or of

tion of associations, *Psychol. Rev.*, 1919, **26**, 465–473. For comprehensive reviews of the work on time interval, see H. M. Wolfle, Conditioning as a function of the interval between the conditioned and the original stimulus, *J. Gener. Psychol.*, 1932, **7**, 80–103; and E. R. Guthrie, Association as a function of time interval, *Psychol. Rev.*, 1933, **40**, 355–367.

[16] E. R. Guthrie, *op. cit.*

[17] H. A. Carr, Teaching and learning, *J. Genet. Psychol.*, 1930, **37**, 189–218.

[18] This factor has already received mention; see W. S. Hunter, Basic phenomena in learning, *J. Gener. Psychol.*, 1933, **8**, 299–317.

range of sensory consequences regarding the error, and other factors, all of which may be grouped under the contemporary conception of effect. Effect depends upon two factors, frequency and temporal extent; and temporal extent, in turn, depends upon frequency. It may be that there is some, as yet unknown, optimal combination of frequency and temporal magnitude.

There is a fundamental difference between the time interval separating items of a list of nonsense syllables or words (or the two stimuli in the conditioned reflex set-up) and the time interval spent in the blind alley (which interval also separates items to be associated, *i.e.*, consecutive parts of the true path). In the latter case, there is a direct relation between length of time interval and thwarting effect, which does not often appear in the former. Thus, lengthening the amount of time in the blind alley has an effect similar to increasing the frequency of the same blind alley. According to this tentative explanation, both time and frequency are viewed as limiting the operation of the law of effect.

Summary

Thirty college students have learned the Warden 10-alley maze to a criterion of three perfect trials in succession. Besides the usual maze data, records have been taken for each subject of the exact amount of time in fifths of a second spent in each entrance into a cul-de-sac. These data were investigated in order to determine the relationship existing between time spent in the cul-de-sac and ease of elimination and certain other time relationships which obtain within the maze pattern. The following conclusions have been obtained:

(1) An inverse relationship exists between the length of time spent in a cul-de-sac (the range of the average entrance times is from one to fifteen seconds) and the difficulty of its elimination.

(2) The percentage of time spent in the true path and in the blind alleys is, in terms of the average, for all subjects, no matter what the time spent in the total learning, nearly a constant.

(3) Serial position effects operate in the learning of a skilled act, but they are not of the same nature as those operating in the memorizing of lists of verbal material. Although there is a general primacy-finality effect, as shown by the average values for the first, last, and middle alleys, the effect is not a regular one. The serial position curve is highly irregular.

(4) The obtained inverse relation between length of time spent in a blind alley and the speed of its elimination may be explained in terms of the greater amount of detrimental sensory consequences which the longer time provides. A factor of *temporal extent* must be added to those influencing learning.

(Manuscript received May 21, 1934)

THE DEVELOPMENT OF THE REACTIVE PROCESS FROM EARLY CHILDHOOD TO MATURITY

BY FLORENCE L. GOODENOUGH

Institute of Child Welfare University of Minnesota

Except for the pioneer study by Gilbert in 1894,[1] few attempts have been made to study developmental changes in simple reaction-time. Nevertheless the problem is one of considerable theoretical interest since the speed with which an individual is able to make a simple motor response to a perceived stimulus may fairly be regarded as a significant index to his basic level of perceptual-motor integration.

In Part III of his recent book, Luria [2] describes a number of interesting experiments on the development of the reactive process in young children but does not give quantitative results. Luria emphasizes the diffuse character of the child's early responses to a signal, his inability to make a single prompt response and to inhibit inadequate additional responses. This ability, he states, is relatively late in appearing. "The reaction to a signal, as we know it in the human adult, is a product of very complex development, an elaboration which arises on the basis of other, considerably more primitive processes. The 'simple reaction' in young children differs from the reaction of adults in having another structure, and is characterized by a marked specificity of the diffused excitation, a weakness of those higher regulating mechanisms which are undoubtedly a basic phenomenon in the neurodynamics of the adult. The development of the reactive process from child to adult does not by any means take place by the quantitative improvement of the process but through a qualitative

[1] Gilbert, J. A., Researches on the mental and physical development of school children, *Studies from the Yale Psychological Laboratory*, 1894, **2**, 40–100.

[2] Luria, A. R., *The Nature of Human Conflicts or Emotion, Conflict and Will. An Objective Study of the Disorganization and Control of Human Behavior.* Translated from the Russian and edited by W. Horsley Gantt with a foreword by Adolf Meyer. New York: Liveright, 1932.

change in structure overcoming primordial diffusiveness and passing over into a new, controlling, intricate, functional, organized structure of the reaction. These simple movements which we speak of as 'simple reactions' of the adult are, really, very late formations, built up on the basis of the suppressed, diffused system of the primitive impulses."[3]

That certain qualitative differences between child and adult in their characteristic manner of responding to a signal indubitably exist will be shown in a later section of this paper. These qualitative differences, however, are not independent of the quantitative aspects of the reaction but have a definite bearing upon the time relationships displayed.

The present study is an attempt to ascertain the changes with age in the speed of a simple reaction during early childhood and the relation of reaction-time at these ages to such factors as sex, intelligence test scores, socio-economic status and certain personality traits.

Reaction-time was measured by means of the Miles Reaction-time Board which has been described elsewhere.[4] This instrument consists of an electrical timing clock which operates on a 60-cycle current in the same manner as an ordinary electric clock and which records time in units of 1/120 of a second. When running, the clock makes a fairly loud buzzing noise. It is started and stopped either by a telegraph key which is wired in circuit or by an accessory push-button switch on an extension cord which the experimenter holds in his hand. Both clock and key are mounted on a light-weight board which makes the apparatus easily portable. The complete apparatus includes an accessory key with extension cord which makes it possible to place the subject at another table out of sight of the clock, in which case he must react to the sound of the buzzer alone. This key was not used in the present experiment.

Because of the youth of our subjects, it was found necessary to modify the instructions used by Miles in certain respects in order to make the procedure clear. Since we also wished to compare the speed of the right and left hands and it was not feasible to recall the children for the extended practice which would have been necessary to establish a true physiological speed limit[5] we have made use of alternating trials with the two hands in order to equalize the effects of practice.

---

[3] *Op. cit.*, p. 338.

[4] Miles, Walter R., Correlation of reaction and coördination speed with age in adults, *Amer. J. Psychol.*, 1931, 43, 337-391.

[5] Even had time and opportunity permitted, it is doubtful whether this could be done with the majority of young children because of the flagging interest and attention which is the almost inevitable accompaniment of prolonged work at a monotonous task with children too young to understand what it is all about. In this case, the novelty of the instrument and the fun of 'stopping the clock' provided a good source of motivation at the outset. By dint of lavish praise and much display of enthusiasm, the ex-

Our procedure was as follows: The board without the accessory key was placed on a low table and connected to an ordinary lighting socket. With the subject seated comfortably before the table the examiner said:

"I have a clock here (pointing) that makes a loud noise and its hands go round just like your clock at home. You can stop the clock if you put your finger down *hard* on this button like this. (Experimenter demonstrates by starting the clock, using the push button, and stopping it by pressing the key. This is repeated, and child is allowed to examine the instrument until he feels entirely familiar with it.) Then the experimenter says, "Now let's see how quickly you can stop it. Put your finger on the button so you will be all ready and as soon as I start it, you stop it. Don't take your finger off till the noise stops, will you?"

After the child had made one trial, the experimenter said, "Now see how quickly you can stop it with the other hand." At least two preliminary trials with each hand were given in this way. With the youngest children it was sometimes necessary to give a little more preliminary practice if the child did not seem to understand that he was to stop the clock as quickly as possible. A few were somewhat frightened by the buzzing sound at first and had to be habituated to it before a prompt reaction could be secured. Others were so delighted by the noise and the movement of the hands that they did not want to stop it. Their interest in 'seeing it go' had to be satisfied before they could be persuaded to devote their attention to 'making it stop.' For this reason it seemed best to allow the amount of preliminary practice to vary somewhat with the personality of the child. The principle followed was to require at least two 'correct' reactions with each hand before starting to record results. A 'correct' reaction was defined as one in which the child evidently tried to stop the clock as soon as it was started.

In the preliminary practice the child was allowed to begin with whichever hand he preferred. The regular series, however, was always begun with the right hand. The experimenter said, "Now that you know how to do it, I'm going to keep track and see how quickly you can stop the clock. Let's get all ready" (placing the forefinger of the child's right hand on the key while his left hand rested easily on the table. "Are you all ready now? Then see how quickly you can stop it."

Care was taken to see that the child was seated comfortably with both feet on the floor. Posture was watched and corrected as necessary throughout the experiment. Before starting the experimental series, the readings for both dials were noted on the record sheet. After each reaction the new reading was taken and the child was praised for his quickness. Alternate trials with each hand were made until 20 trials, 10 with each hand had been completed.

---

perimenter was able to prolong the child's interest to the end of the single series of 20 trials in practically all cases. It was found, however, that when a second sitting was attempted the maintenance of a high degree of effort on the part of the younger children was very difficult to secure, and even those of six and seven years were inclined to attack the task with less zest. As a result, whatever gain might have been expected to come from the increased practice was so masked by the results of decreasing interest that the reaction-time on the second day of practice was in most cases no quicker and with the youngest children might even be slower than that which had been established on the first day. Our procedure is similar to that followed by Miles and by Gilbert neither of whom put their subjects through the period of preliminary training recommended by Titchener and others. Miles refers to the reaction-time of unpracticed subjects as the 'natural' reaction-time.

Our data consist of 466 complete series of 20 trials each from 246 children ranging in age from two-and-a-half [6] to eleven-and-a-half years and 56 college students. Of these, 106 children were retested after an interval of one year and 58 after an interval of two years. Tests were always given midway between birthdays. The maximum variation from the defined age-level was one month in either direction, *e.g.*, the five-and-a-half-year age-group includes only children between the ages of five years, five months and five years, seven months, etc. The subjects include children enrolled in the nursery school and kindergarten of the University of Minnesota Institute of Child Welfare and others who are brought to the laboratory for periodic tests. These children are all well habituated to the test situation and their cooperation is in most cases very easy to secure.

The question of scoring presented some difficulties. In reaction-time work with adults it is customary to make use of introspective reports by the subjects as a criterion for eliminating trials in which, owing to temporary distractions or other causes, the reaction is unduly prolonged. Obviously, such reports would be of little value when given by young children. Gilbert, when faced by the same problem, hit upon the expedient of using the median of all the trials in place of the mean as the most typical measure of the reaction-speed of an individual child. As a check upon the desirability of this procedure we have computed the standard error of each of these measures for the right and left hands separately at each of three age-levels. The results are shown in Table 1.

Inspection of Table 1 shows that the standard error of the median is less than that of the mean in every one of the six comparisons made. We have therefore used the median in preference to the mean as the measure of individual reaction-time throughout the study.

[6] The six children who were tested at age two-and-a-half, cannot, however, be looked upon as representative of two-and-a-half-year-old children in general since only a small proportion of children at that age were found able to understand the instructions and give sustained attention to the task for the requisite number of trials. We have therefore omitted these cases from all computations into which the question of normative standards enters. From the age of three-and-a-half on, practically all children of normal intelligence were able to take the test.

TABLE 1

COMPARATIVE RELIABILITY, AS INDICATED BY THEIR STANDARD ERRORS, OF MEASUREMENTS BASED UPON THE MEAN AND THOSE BASED UPON THE MEDIAN

	(In sigma units)					
Age.....	3½ (N = 53)		5½ (N = 111)		7½ (N = 53)	
	Right	Left	Right	Left	Right	Left
σ mean.....	3.62	3.19	1.30	1.50	.85	.90
σ median.....	3.00	2.81	1.20	1.32	.72	.75

Table 2 shows the median reaction-time in sigma units for right and left hands separately for the three ages included in Table 1 and for the college students. Surprisingly enough, the reaction-time of the two hands seems to be practically identical at these ages, under the conditions of this experiment.

TABLE 2

MEDIAN REACTION-TIME OF RIGHT AND LEFT HANDS AT FOUR AGE-LEVELS

Age.....	3½			5½			7½			College Students		
Hand	Boys	Girls	Both	Boys	Girls	Both	Boys	Girls	Both	Men	Women	Both
Right...	473	510	491	325	348	336	243	253	248	173	171	172
Left....	494	527	510	302	340	320	249	248	249	169	173	172
Both....	484	519	501	313	345	329	246	251	249	171	172	172

The same is true when the mean instead of the median is used as the measure of individual reaction-time. Inspection of the records of individual cases showed little consistent superiority of either hand from trial to trial. For this reason it has not seemed worth while to treat the results from the two hands separately.

Table 3 shows the median reaction-time by ages and sexes separately, together with the means and the quartile deviations. This table also shows the results obtained from 56 Minnesota college students who were tested by the same method as that used with the children. Miles [7] findings for older adults are shown in Table 4.

[7] *Op. cit.*

The Minnesota students were tested by exactly the same procedure as was used with the children and were given the same number of trials. Miles' procedure, however, was slightly different in that the 'right or preferred hand' was used throughout and only five trials were given, apparently without the preliminary trials which we used with children and adults alike. Miles' results are given only in terms of mean reaction-time. However, the results presented in Table 3 suggest that the difference betwen the mean and the median is very small after the age of ten or thereabouts.[8]

In comparing these results with those of other investigators it must be remembered that only enough preliminary practice to make sure that the instructions had been thoroughly understood was given. Although the stimulus was both auditory and visual, it seemed fairly certain that the auditory stimulus was practically always the one to which the child responded. This is, of course, to be expected, since it is well known that the reaction to an auditory stimulus is appreciably quicker than that to a visual stimulus. Throughout the experiment it was noted that the behavior and postures of the children were characteristic of 'listening' rather than 'watching.' The presence of the clock added interest to the game and after a reaction had been made the child usually looked to make sure that the hands had stopped, but during the moment of waiting for the signal his eyes were much more likely to be fixed on the experimenter than on the clock. It is probable, however, that in some cases the attention was centered on the visual stimulus and to the extent that this was the case the reaction-times would be somewhat lengthened. From our observations of the children, however, we are inclined to believe that these cases represent a small proportion of the total. When the

[8] This is in part due to the fact that we made no use of introspective reports either with children or adults. With the former it would obviously have been impossible to secure them. Although long delays resulting from an obvious distraction as well as reactions so short that it seemed certain that the subject had 'jumped the gun' were omitted, nevertheless a good deal more leeway in both directions was allowed than has been done by many investigators in the past. The result of this procedure has been to make the distribution of the individual reaction-times for the older children and adults conform fairly closely to the normal probability curve, while those for the younger children show marked skewness.

TABLE 3

MEANS, MEDIANS, AND QUARTILE DEVIATIONS OF REACTION TIME BY AGE AND SEX

(In sigma units)

	Age																		
	3½		4½		5½		6½		7½		8½		9½		10½*	11½*	College Students		
	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.	M.	F.	Both	Both	M.	F.	Both
Mean	580	622	406	458	364	414	282	309	279	260	230	262	244	205	229	198	174	176	175
Median	492	518	356	424	311	356	259	286	260	250	223	249	218	202	229	192	171	172	172
75 percentile	324	372	259	327	238	265	211	228	210	208	164	204	159	149	173	144	147	148	148
25 percentile	673	763	481	543	429	480	322	370	323	292	272	293	276	257	274	250	187	190	189
Av. m.v. from median	226	236	116	127	106	127	66	73	63	51	44	59	46	32	36	31	25	26	26
Av. m.v. from mean	232	247	133	130	112	133	70	76	66	51	46	63	48	36	38	32	26	27	27
No. Cases	26	27	28	29	54	57	25	34	24	29	18	12	13	9	12	7	13	43	56

* Because of the small number of cases, sexes have been combined at ages 10½ and 11½.

curve obtained for the children is extrapolated, it fits the average for the college students very well. It appears, therefore, that the percentage reacting to the visual rather than to the auditory stimulus when both are presented simultaneously shows little change with age. Miles' results, however, tend to run somewhat higher than those for our college students. This may be due to the small number of cases in each of his groups or to the fact that he gave only five trials without fore-exercise.

TABLE 4

MEAN TIMES OF REACTION* BY PRESSING THE FINGER FOR ADULT SUBJECTS OF VARIOUS AGES AS REPORTED BY MILES

	Age-groups								
	25–29	30–34	35–39	40–44	45–49	50–59	60–69	70–79	80–89
Mean	230	240	230	210	230	200	280	300	280
Av. m.v.	30	80	40	30	40	30	110	80	30

* For the 74 cases under the age of 60 the mean reaction-time is 223; the mean variability from trial to trial is 42. For the 26 cases aged 60 or over, the mean reaction-time is 287 and the variability 73. Expressing these figures in terms of their reciprocals, that is in terms of speed and non-variability, it is seen that the older individuals have, on the average, suffered a decrement of 22.2 percent in speed of reaction and 42.5 percent in regularization or control of the reaction. In other words, the age decrement in control is almost double the corresponding decrement in average speed.

At these early ages it would obviously be very difficult to train the subjects to differentiate between the sensorial and the muscular types of response and we made no attempt to do so. Titchener, in his *Textbook*, gives the figures shown in Table 5

TABLE 5

SENSORIAL AND MUSCULAR REACTION-TIMES TO LIGHT AND TO SOUND
(From Titchener)

	Muscular	Sensorial
Light	180	290
Sound	120	225

as averages for the two extreme types of reaction-times with moderately practiced subjects. He states that for the mixed reaction in which the attention is sometimes centered on the movement, sometimes on the stimulus, the average times will fall somewhere between the two extremes cited in Table 5.

Ladd and Woodworth summarize the determinations of many of the older investigators. For optical stimuli the reported times vary from 151 (Exner) to 225 (Hankel). The average for the eight investigators whose results are included in the Ladd and Woodworth table is 190. The corresponding average for auditory stimuli is 146 with a range from 120 (Von Kries), to 182 (Von Wittich). All these figures are based upon adults. Our results for college students fall well within Titchener's range for the mixed reaction to sound as well as within the range reported by other investigators who have not distinguished between the sensorial and the muscular types of reaction.

Inspection of Table 3 shows that for the ages covered by this study there is a fairly regular decrease in reaction-time with age and that the boys, on the average, react somewhat more quickly than the girls. The sex difference corresponds very closely both in direction and amount to that reported by Gilbert [9] who used a visual stimulus and whose study covers an age-range from six to seventeen years. Bellis [10] who tested 10 subjects of each sex in each age-decade from four to sixty years reports a very marked sex difference in the time of reaction both to light and to sound but these differences are spuriously high because of his failure to equalize age within the decades. Among the children and young adults the girls in a given age-group have an average age much younger than that of the boys; among the older adults the women average somewhat older than the men. If his data are re-plotted on a comparable age-basis, however, the sex difference in favor of the males is still apparent though it is much smaller than that reported by Bellis. Although this difference is not great at any age, the fact that it appears in children as young as three years and is maintained without appreciable change at least until the upper fifties is highly significant and well merits further study. The superiority of the males is shown even more clearly in Table 6 which shows the percentages of the individual reactions which were delayed beyond certain

[9] *Op. cit.*

[10] Bellis, C. J., Reaction time and chronological age, *Proc. Soc. Exper. Biol. and Med.*, 1933, **30**, 801–803.

specified times. At all ages except seven-and-a-half, the percentages of delayed reactions are distinctly greater for the girls than for the boys. After the age of nine-and-a-half, however, the proportion of delayed reactions becomes very small for both sexes and among the college students, reactions delayed for as long as .5 second are rare.

Not only do the median reaction-times decrease regularly with age but the spread of the individual reactions about the individual medians shows an even more rapid decrease with age. This is shown in the lower part of Table 3 which presents the means of the individual mean variations about the individual central tendencies. In order to make these figures correspond more closely with those which have been reported in the literature, we have calculated the m.v.'s in two ways; first about the medians as central tendencies (which obviously gives the minimum figure) and secondly in the usual manner reckoning the variations from the individual means.

TABLE 6

PERCENTAGES OF INDIVIDUAL REACTION-TIMES DELAYED BEYOND CERTAIN LIMITS

Percent of all reactions which were longer than	Age												
	3½		4½		5½		6½		7½		8½		9½–11½
	B.	G.	B.	G.	B.	G.	B.	G.	B.	G.	B.	G.	Both
0.5 sec....	48.7	52.9	21.9	30.9	14.9	21.5	5.5	6.6	6.4	2.7	2.2	3.3	1.0
1.0 "	12.4	13.2	2.6	2.7	1.7	3.8	0.2	0.3	0.0	0.0	0.0	0.0	0.1
2.0 "	1.3	2.1	0.5	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Examination of Tables 3 and 6 shows that there are two distinct types of developmental change in the ability to react to a signal. Not only do children react more quickly, on the average, as they grow older, but their reactions become more stable with fewer marked fluctuations in speed from one trial to another. Observation of the behavior of the children leaves little doubt that in the majority of instances, at least, the extreme fluctuations in reaction-speed which are so characteristic of the younger ages cannot fairly be ascribed to distraction from external sources. Rather they seem to be

indications of what Luria describes as a state of diffused nervous excitation during which the child is truly unable to make the single integrated movement necessary to press the key, even though all his attention is apparently centered upon doing so. Often, at such times, the increased tension of the muscles, not only of the arm and hand but of the entire body, is so great as to be plainly discernable to the experimenter. The child's body stiffens, his eyes are fixed on the apparatus, the breath may be held for an instant and expelled in a sigh of relief when finally the reaction is accomplished. Accessory movements of the resting hand and even of the feet and legs are common. Although Luria gives no quantitative results, he states that in the normal child the primitive diffused reaction has pretty well disappeared by the age of seven or eight years and has been replaced by the more constant integrated reaction which is characteristic of the adult. Our own findings are in agreement with this. After the age of seven-and-a-half, reaction-times greater than half a second rarely occur and the useless accompanying movements of other parts of the body have for the most part disappeared. Although the occasional long reactions which are still to be found in the records of the older children may represent temporary reversions to a more primitive reaction type, it seems likely that in most cases they are the result of true distractions which the experimenter was unable to detect.[11]

It is interesting to note that the changes in the reactive process during old age which have been reported by Miles show a corresponding trend. (See Table 4.) Here, too, the change in variability is proportionately much greater than the change in speed. At the age of three-and-a-half, the average child has attained 34.9 per cent of adult speed of reaction but only 11.2 percent of adult regularity in the control of the reaction. According to Miles' figures, adults of 60–89 years react 77.8 percent as speedily as those under the age of

[11] Great care was taken to make sure that the child was always 'set' to respond before the signal was given. If, in spite of this, there was reason to think that a distraction had occurred, the reaction was not recorded. However, since the detection of all distracting influences is extremely difficult to accomplish, it is probable that an occasional instance of this sort has crept into the results.

60 but regularity of control has dropped to 57.5 percent of its former level.

Individual differences among the subjects are marked at all age levels. This is true both of the median speed of reaction and of the mean variations from the individual medians. The reliability of the medians has been computed by the 'split-scale' method and the results thus obtained have been corrected to correspond to the true length of the test series by means of the Spearman-Brown formula. The results, which are presented in Table 7, show a fairly satisfactory degree of

TABLE 7

ODD-EVEN RELIABILITY COEFFICIENTS FOR INDIVIDUAL MEDIANS COMPUTED BY THE PEARSON PRODUCT-MOMENT METHOD OF CORRELATION AND CORRECTED BY THE SPEARMAN-BROWN FORMULA

Age	Boys		Girls		Sexes Combined	
	N	*r*	*N*	*r*	*N*	*r*
3½	26	.853	27	.759	53	.810
4½	28	.782	29	.755	57	.784
5½	54	.903	57	.900	111	.894
6½	25	.859	34	.825	59	.847
7½	24	.897	29	.903	53	.900
8½	18	.885	12	.886	30	.880
9½	13	.921	9	.873	22	.905
10½–11½	—	—	—	—	19	.802
College students	13	.726	43	.904	56	.886

reliability at all ages with a slight tendency toward increase in the magnitude of the coefficients as age advances. In five of eight comparisons the reliability is slightly higher for the boys than for the girls. While this difference is well within the limits of chance, the tendency is in line with the other evidence pointing toward a small degree of male superiority in reaction-speed, since reliability as well as speed shows some tendency to improve with age.

Table 8 shows the corresponding reliability coefficients for the mean variations as computed from the individual medians. These have been computed for three age-groups only. While these coefficients are somewhat lower than those presented in Table 5, they are nevertheless high enough to indicate that

variability in reaction-time as well as its speed is a significant measure of the individual.

We were interested in seeing to what extent individual differences in reaction-speed were maintained over a period of time. As was stated before, 106 children were retested after a period of one year and 58 after an interval of two years. The correlation between the individual medians after a one-

TABLE 8

RELIABILITY OF THE MEAN VARIATIONS FROM THE MEDIAN CALCULATED BY THE PRODUCT-MOMENT METHOD AND CORRECTED BY THE SPEARMAN-BROWN PROPHECY FORMULA

Age	3½		5½		7½	
	r	*N*	*r*	*N*	*r*	*N*
Boys	.742	26	.693	54	.693	24
Girls	.823	27	.752	57	.648	29

year interval was .675 which becomes .469 when age is held constant by partial correlation. As a check on the partial correlation technique we computed the correlation separately for a group of 37 children all of whom were aged five-and-a-half at the time of the first test and six-and-a-half at the time of the second test. This correlation was found to be .443. The difference is well within the limits of chance variation. For the 58 cases retested after an interval of two years, the correlation was .540 which becomes .370 with age held constant.

In like manner, the correlations between the mean variations from the individual medians were computed for the same subjects. For the single-year interval the correlation was found to be .454 which becomes .283 with age held constant. After a two-year interval the correlation is only .216 and this becomes practically zero when age is partialled out. The lower reliability of the mean variations must be considered in interpreting these results.

Whether or not individual differences in reaction-speed are significantly related to general intelligence or 'brightness' is a question that has attracted the interest of psychologists since the days of Gilbert. Gilbert, working before the days of

intelligence tests, made use of teachers' judgments of the ability of their pupils by way of throwing light on this problem The children were classified into three groups, bright, average, and dull. When the median reaction-times of each group were plotted according to age, a very slight and irregular tendency for the bright to exceed the average and the average to exceed the dull in reaction-speed was found to exist. Nevertheless, the differences are so small and the curves cross and recross at so many points that only by careful inspection of the results for the entire range of ages can the general trend be detected. The correlation is evidently positive but very low.

A somewhat more recent study by Peak and Boring [12] which was based on a comparison of the reaction-speeds of five college students with their scores on the Otis Intelligence Test and the Army Alpha Test under various conditions of timing has been much quoted. Their results appeared to show that, in the words of the authors, "there is a high correlation between score in an intelligence test, speed in an intelligence test, and speed in a simple reaction." However, because of the small number of subjects and the fact that these results were so much out of line with those obtained by previous workers, a check study using the same procedure was made by Farnsworth, Seashore and Tinker [13] with 35 college students as subjects. Their results do not confirm those of Peak and Boring. Practically a zero relationship was found between the intelligence tests used and simple reaction-time. The relationship between the tests and serial reaction-time was somewhat higher than that found for the simple reaction.

Beck,[14] using as subjects 30 college students, found very low though generally positive correlations between speed of reaction to a visual stimulus and scores on the Thorndike Intelligence test and on rate of reading. The reliability of the

[12] Peak, H. and Boring, E. G., The factor of speed in intelligence, J. Exper. Psychol., 1926, **9**, 71–94.

[13] Farnsworth, P. R., Seashore, R. H. and Tinker, M. A., Speed in simple and serial action as related to performance in certain 'intelligence' tests, *Ped. Sem.*, 1927, **34**, 537–551.

[14] Beck, L. F., Relation of speed of reaction to intelligence. *Amer. J. Psychol.*, 1932, **44**, 793–795.

reaction-time scores for 100 trials given in 4 series of 25 trials each was found to be +.93.

From these and other studies, it may be concluded that among adults, at least, there is little relationship between simple reaction-speed and intelligence level when the subjects are sufficiently homogeneous with respect to ability for an intelligence test to be useful as a means of distinguishing between them. One does not need a test to tell the genius from the idiot, and in most cases, everyday observation will soon enable one to separate the extremes of ability even at the college level. The 'high correlation' reported by Peak and Boring is apparently found only among subjects of widely variant abilities. Among groups selected at random from a reasonably homogeneous population, the relationship is very slight.

Nevertheless, because of the fact that even among children of the same age, significant relationships between traits have sometimes been found to exist during the developmental period which either disappear or become much lower after maturity has been reached, it seemed worth while to compare the findings from the reaction-time tests with the results of a series of intelligence tests given to each child within a few days of the date of the reaction-time experiment. These tests included, at different age-levels, the Minnesota Preschool Tests, the Merrill-Palmer Tests, the Arthur Performance Scale and the Stanford Binet. No cases in which the reaction-time test was separated from the intelligence test by more than a two-week interval have been included. The relatively small number of cases who were given the Stanford is due to the fact that in a large number of the cases at the older ages, a preliminary form of the new revision of the Stanford which is as yet unpublished was used. Since data for combining the results of the old and the new revisions are not yet available we have been obliged to omit the cases who were given the new form in calculating the correlations presented in Table 9. The Minnesota Tests are divided into two parts, a verbal scale and a non-verbal scale, which are scored both separately and in combination. Both the verbal and the non-verbal scales are

scored wholly in terms of quality of performance; time is not a factor. The Merrill-Palmer Tests and the Arthur Scale are chiefly non-verbal (except for the matter of comprehending instructions the Arthur is wholly so), and a large percentage of the individual test items are scored in terms of time. Although a very small number of the items of the Stanford Binet are time-limited, the time allowances are so generous that practically all children who are capable of performing the tasks have ample time to do so. It is fair to say, therefore, that while speed of performance is a partial determinant of scores on the Arthur and the Merrill-Palmer tests, the Minnesota and the Stanford scores are independent of speed.

Table 9 shows the correlations by ages and sexes separately between reaction-speed and each of the two tests in which

TABLE 9

CORRELATIONS * BETWEEN REACTION-SPEED AND SCORES ON INTELLIGENCE TESTS NOT INVOLVING SPEED

	Age					
	3½		4½		5½	
	B	*G*	*B*	*G*	*B*	*G*
Minn. Verbal	+.01	+.19	−.31	−.19	+.33	+.04
Minn. Non-Verbal	+.21	+.30	+.26	+.13	+.14	+.07
Combined Scale	+.18	+.26	−.11	−.05	+.29	+.08
N	22	26	23	28	52	56
	5½		6½		7½	
	B	*G*	*B*	*G*	*B*	*G*
Stanford-Binet	+.43	+.57	+.11	−.12	+.22	−.17
N	21	12	13	14	15	17
	8½		9½		10½–11½	
	B	*G*	*B*	*G*	*B*	*G*
Stanford-Binet	+.10	−.03	−.05	−.22	−.64	+.39
N	8	7	11	8	7	6

* The Pearson product-moment formula was used in computing the correlations with the Minnesota test, the Spearman rank-order method for the Stanford.

speed is not a factor. Table 10 shows the corresponding correlations for the tests of which the scores are in part determined by speed.

TABLE 10

CORRELATIONS * BETWEEN REACTION-SPEED AND SCORES ON INTELLIGENCE TESTS IN WHICH SPEED IS A FACTOR

	Age					
	3½		4½			
	B	*G*	*B*	*G*		
Merrill-Palmer *r*	+.27	+.41	+.34	+.33		
N	22	23	23	26		
	5½		6½		7½	
	B	*G*	*B*	*G*	*B*	*G*
Arthur *r*	+.35	+.08	+.39	+.24	+.58	+.31
N	50	47	13	22	17	22
	8½		9½		10½–11½	
Arthur *r*	+.04		+.12		+.16	
N	20		21		13	

* Because of the small number of cases who were given the Arthur tests after the age of 8½ the sexes have been combined for these correlations. The Pearson product-moment method was used.

The correlations shown in Table 9 are low and, in view of the small number of cases at each age, are well within the limits of chance variation from zero. However, the fact that 20 of the 30 coefficients included in this table are positive in sign, suggests that a more extensive study would probably reveal a very low positive relationship between reaction-speed and scores on intelligence tests that do not involve speed. There is no indication, however, that this relationship is any closer during childhood than that previously reported for the adult level.

For the tests included in Table 10, the relationship is somewhat closer, though it is still rather low. All the coefficients in this table are positive. The comparatively low correlation between reaction-time and scores on the Arthur

test, particularly at the upper ages is probably due, at least in part, to the low reliability of this test.[15] Kawin [16] has recently reported a reliability coefficient of +.59 for the Merrill-Palmer Scale. It is reasonable to suppose, therefore, that the correlations which were obtained between these tests and reaction-speed have been 'attenuated' by the unreliability of the measuring scales used to a much greater extent than those presented in Table 9 [17] and that if the 'speed tests' used had been equally as reliable as the 'power tests,' the difference between their respective correlations with the reaction-time scores would have been greater than that obtained in this study. Because of the small number of cases included in each age-group it did not seem advisable to correct for the attenuation factor since the probable error of such a correction would be very large. The direction of the change that would have been introduced is, however, fairly certain, though its amount is undetermined.

Comparison of the mean reaction-speed of children from the upper socio-economic classes with that of the children from the lower social strata revealed no consistent differences between the groups. This is in line with other studies which have commonly shown little difference between social classes in motor skills although tests of abstract intelligence and of facility in the use of language commonly show a marked superiority of the more favored socio-economic groups.

Correlations of reaction-speed with height and with weight also tend to be positive but very low. With age held constant, the correlation between speed of reaction and standing height was found to be +.061 for boys and +.128 for girls. The corresponding figures for weight were +.248 for boys and +.044 for girls.

[15] See Table 29, pp. 69–79 in *A Point Scale of Performance Tests*, Vol. II, by Grace Arthur. From this table it would appear that the correlation between Form I and Form II of the Arthur test averages not over .60 for a single-year age-group and that its reliability decreases with age.

[16] Ethel Kawin, *Children of Preschool Age*, Chicago: University of Chicago Press, 1933.

[17] The reliability of both the Stanford Binet and the Minnesota Preschool Scale has been found to be in the neighborhood of .90 for representative groups of the same age.

Observational records of certain forms of personal-social behavior obtained by the time-sampling method were made of all children in the nursery school and kindergarten at the time that the tests were given. On the hypothesis that children who are physically less active than the average may show a corresponding retardation in reaction speed, we correlated their physical activity scores (based on the average of 40 half-minute samples taken on 20 successive days) with their median reaction-times, keeping ages and sexes separate. Five of the six correlations are positive (greater activity going with shorter reaction-time), and one slightly negative. The values range from −.09 to +.20 for groups varying in size from 10 to 36 cases each.

A somewhat higher relationship seems to exist between speed of reaction and frequency of laughter, for which all six correlations are positive with values ranging from slightly above zero to +.52. For the two largest groups, consisting of 34 five-and-a-half year old boys and 36 five-and-a-half year old girls the correlations with physical activity are +.201 and +.169 respectively while those with laughter are +.300 and +.194. While these findings need to be confirmed by further studies with larger groups, they raise a number of interesting theoretical questions as to the role of the reactive process in the organization and integration of the personality.

Summary

1. By the use of the Miles Reaction Board it was found possible to determine the speed of simple reaction to a predominantly auditory stimulus with a satisfactory degree of reliability in children as young as three-and-a-half years.

2. The development of the reactive process during childhood is shown not merely by improved speed of reaction but to an even more marked degree in the gaining of voluntary control over the motor act. The improvement in voluntary control is shown in the gradual reduction of useless accessory movements preceding and accompanying the act of pressing the reaction-key, and in fewer signs of bodily tension as age advances. A quantitative indication of this aspect of de-

velopment is given by the changes with age in the variability in the speed of successive reactions by the same individual. The developmental changes in variability are much more marked than the improvement in average speed. Examination of data collected by Miles suggests that a similar trend can be noted during the years of mental decline. Both the age-increment during childhood and youth and the age-decrement during senescence appear to be more clearly shown in the variability of the reaction than in its average speed.

3. The slight sex difference in favor of the males which previous investigators have reported for adult subjects appears to hold good even in early childhood. The boys in our group tend to surpass the girls of the same age both in respect to average speed of reaction and in low variability from trial to trial. Sex differences, however, are very small in amount and there is much overlapping between the sexes.

4. In agreement with the usual findings for adults, only a slight relationship was found between scores on intelligence tests which do not involve speed and the speed of a simple reaction for representative groups of child subjects. The correlations between speed of reaction and performance tests in which speed is a factor, although not high, were positive at all ages and for both sexes. No relationship between speed of reaction and socio-economic status was apparent. The relationship of reaction-speed to height and weight was positive but very low.

5. The degree of physical activity as indicated by time-sampling observations in the nursery school and kindergarten showed a small positive relationship to speed of reaction in 5 of 6 groups studied. Amount and frequency of laughter during free play was found to be positively related to reaction-speed in all of the six groups studied but the correlations were not high.

(Manuscript received June 6, 1934)

AN ATMOSPHERE EFFECT IN FORMAL SYLLOGISTIC REASONING

BY ROBERT S. WOODWORTH AND SAUL B. SELLS

Columbia University

True though it certainly is that the syllogism is a tool of logical analysis rather than a diagram of any typical reasoning process, the use of syllogisms as problems appears to be a promising lead in the experimental study of thinking. Especially syllogisms with purely symbolic terms, such as

All S is M;
All P is M;
Therefore, all S is P,

have some of the advantages possessed by lists of nonsense syllables in memory experiments, in their freedom from extraneous associations and from factual truth or falsity. When such symbolic syllogisms, valid and invalid, are presented in an experiment, as has been done by Störring (**2–7**) and by Wilkins (**8**), they are found to be very unequal in difficulty. Certain of the standard logical fallacies, like that of Undistributed Middle, offer much more difficulty than some others, as Universal Conclusion from Particular Premise. Why some syllogisms make easy problems and others difficult ones is a question for psychological inquiry.

In scrutinizing the detailed reports of Störring and Wilkins, we have succeeded in formulating three hypotheses which, taken together, seem to account for the observed differences in difficulty, the presence of difficulty being indicated either by the subject's introspective report or by the objective fact of his acceptance of an invalid conclusion.

The first hypothesis will be readily accepted. It is simply that difficulty arises from the ambiguity of the language in which syllogisms are expressed. Ambiguity attaches especially to the word 'some,' as in "Some *X* is *Y*." By the

conventions of formal logic 'some' means 'at least some,' *i.e.* 'some and perhaps all,' while in ordinary speech it often carries the implication, 'some but not all.' A person not familiar with formal logic may therefore regard it as perfectly correct to infer from "Some *X* is *Y*" that "Some (other) *X* is not *Y*."

The second hypothesis is that of caution or wariness on the part of the subject in an experiment. It is evidently more incautious to accept a universal than a particular conclusion, and probably it is more incautious to accept an affirmative than a negative proposition. At any rate a larger percent of invalid particular conclusions are accepted than of universal, and of negative than of affirmative.

The third hypothesis, that of 'atmosphere,' demands fuller explanation and defence. The atmosphere of the premises may be affirmative or negative, universal or particular. Whatever it is, according to the hypothesis, it creates a sense of validity for the corresponding conclusion. An affirmative atmosphere in the premises makes it easy to accept an affirmative conclusion, etc.

An analogous kind of 'atmosphere' accounts for many grammatical errors. A heedless speaker or writer often makes the verb agree in number with the general atmosphere of the subject phrase instead of with the noun which is the formal subject. For instance:

"The laboratory equipment in these situations were in many instances essentially the same as those used before."

"A common background of emotional conflict, repression, and abnormal reactions of a motor, visceral or perceptual nature, have justified an inherent relationship."

Still another analogy is found in the 'confluxion' illusions, such as the Müller-Lyer, in which the observer sees one of two equal lines as longer than the other because the total impression of one part of the figure is bulkier than that of the other part. He finds it difficult in the figure and in the sentence, as in the syllogism, to analyze out the precise fact or relationship with which he is trying to deal.

The atmosphere hypothesis was first suggested by the difficulties of 'simple conversion,' as brought out by Eidens

(1) and by Wilkins. If we use the customary symbols for the four standard propositions, we have the following table, derived from the data of Wilkins.

PERCENT OF COLLEGE STUDENTS ACCEPTING THE SIMPLE CONVERSE OF SYMBOLIC PROPOSITIONS

A = universal affirmative, as "All *X* is *Y*"	24%
E = universal negative, as "No *X* is *Y*"	72
I = particular affirmative, as "Some *X* is *Y*"	90
O = particular negative, as "Some *X* is not *Y*"	53

'Atmosphere' can account for the acceptance of the simple converse of *A* or *O*, which is invalid, as well, no doubt, as for some of the acceptances of the valid converse of *E* or *I*. 'Caution' can explain the failure of a considerable percent to accept the converse of *E*, and the less frequent acceptance of the converse of *A* than of *I*.

According to the atmosphere hypothesis,

A has an all-yes atmosphere
E has an all-no atmosphere
I has a some-yes atmosphere
O has a some-no atmosphere

When the subject does not *see* the relationships clearly, he is influenced by the atmosphere of the premise, and is inclined to accept the simple converse of any symbolic proposition.

The hypothesis needs to be supplemented or elaborated in order to cover the case of the syllogism with its two premises. If they are both of one kind, *A*, *E*, *I* or *O*, atmosphere clearly calls for a similar conclusion. But if one premise is *A* and the other *I*, the atmosphere is partly universal and partly particular, a blend of *all* and *some*, which would certainly be weaker than a straight *all*, and thus would amount to a *some*.

In general formulation the secondary hypotheses suggested for application to syllogistic reasoning are (1) that a particular premise creates a *some* atmosphere, even though the other premise be universal, and (2) that a negative premise creates a negative atmosphere even though the other premise be affirmative.

In detail, the secondary atmosphere hypotheses are that:
with premises *AA*, atmosphere calls for an *A* conclusion;
with premises *AE* or *EE*, it calls for an *E* conclusion;
with premises *AI* or *II*, it calls for an *I* conclusion;
with premises *AO*, *EI*, *EO*, *IO or OO*, it calls for an *O* conclusion.

The reader can judge for himself whether there is any prima facie case for atmosphere by glancing through the following list of syllogisms, asking himself simply which conclusions seem easiest to accept.

AA If all *X*'s are *Y*'s, and all Z's are *Y*'s, then:
A All *X*'s are Z's
I Some *X*'s are Z's
O Some *X*'s are not Z's
E No *X*'s are Z's

AI If all *X*'s are *Y*'s, and some *Y*'s are Z's, then:
A All *X*'s are Z's
I Some *X*'s are Z's
O Some *X*'s are not Z's
E No *X*'s are Z's

AO If all *X*'s are *Y*'s, and some Z's are not *X*'s, then:
A All Z's are *Y*'s
I Some Z's are *Y*'s
O Some Z's are not *Y*'s
E No Z's are *Y*'s

AE If all *X*'s are *Y*'s, and no *X*'s are Z's, then:
A All Z's are *Y*'s
I Some Z's are *Y*'s
O Some Z's are not *Y*'s
E No Z's are *Y*'s

II If some *X*'s are *Y*'s, and some *Y*'s are Z's, then:
A All *X*'s are Z's
I Some *X*'s are Z's
O Some *X*'s are not Z's
E No *X*'s are Z's

IO If some *X*'s are *Y*'s, and some *Y*'s are not Z's, then:
A All *X*'s are Z's
I Some *X*'s are Z's
O Some *X*'s are not Z's
E No *X*'s are Z's

IE If some *X*'s are *Y*'s, and no *Y*'s are Z's, then:
A All Z's are *X*'s
I Some Z's are *X*'s
O Some Z's are not *X*'s
E No Z's are *X*'s

OO If some *X*'s are not *Y*'s, and some *X*'s are not *Z*'s, then:
- *A* All *Y*'s are *Z*'s
- *I* Some *Y*'s are *Z*'s
- *O* Some *Y*'s are not *Z*'s
- *E* No *Y*'s are *Z*'s

OE If some *X*'s are not *Y*'s, and no *Y*'s are *Z*'s, then:
- *A* All *Z*'s are *X*'s
- *I* Some *Z*'s are *X*'s
- *O* Some *Z*'s are not *X*'s
- *E* No *Z*'s are *X*'s

EE If no *X*'s are *Y*'s, and no *X*'s are *Z*'s, then:
- *A* All *Y*'s are *Z*'s
- *I* Some *Y*'s are *Z*'s
- *O* Some *Y*'s are not *Z*'s
- *E* No *Y*'s are *Z*'s

The more definite evidence which we have to present as a check on the atmosphere hypothesis comes from two syllogism experiments, one being that of Wilkins, already cited, and the other an unpublished study by Sells. Wilkins had as subjects 81 Columbia College undergraduates, while Sells had 90 educated adults who were regularly employed as subjects on the Civil Works Administration project in Adult Learning, conducted at the Institute of Educational Research, Teachers College, Columbia University, under the direction of Dr. Irving Lorge, in 1934.

The syllogisms were presented somewhat differently in the two experiments, and in neither exactly as shown in the above list. Wilkins presented three conclusions with each of 35 pairs of premises, as for example:

No *a*'s are *b*'s; some *a*'s are *c*'s;
therefore

a. some *c*'s are not *b*'s (accepted by 71 percent)
b. no *c*'s are *b*'s (accepted by 23 percent)
c. some *b*'s are not *c*'s (accepted by 47 percent)

Here the first conclusion is valid, and usually in the Wilkins experiment one or more of the suggested conclusions was valid (instead of all being invalid as in our list above). The subjects were instructed to mark each separate conclusion as valid or invalid. The results of this experiment, so far as they have to do with the acceptance of invalid conclusions

from purely symbolic premises, have been extracted by the authors from the Wilkins data and are presented here in Table I.

In the Sells experiment 300 syllogisms were presented, each with a single conclusion to be marked as 'true' or 'false,' as in the example,

If no *x*'s are *y*'s,
And if all *z*'s are *y*'s;
Then no *z*'s are *x*'s.

The results of this experiment by Sells are shown in Table II.

The question answered by Tables I and II is this: Given two premises of the types shown in the left-hand column, and a suggested conclusion of the type shown at the top of a column, what percent of the subjects accept the conclusion as valid when, logically, it is invalid? Each entry shows the responses of the 81–90 subjects to one or more problems; but where there are blanks in the tables, no problems of the corresponding type were presented.

The bold-faced numbers in the tables show the percent accepting invalid conclusions which are favored by 'atmosphere'; and it will be seen that these percents are always substantial, and larger, with few exceptions, than the percents accepting invalid conclusions not favored by atmosphere. The exceptions, and indeed nearly all the acceptances not accounted for by atmosphere, can be brought under the other two hypotheses of 'caution' and of the ambiguity of *some*. For many subjects, "Some are" implies "Some are not" and vice versa. When therefore an *I* conclusion is favored by atmosphere, we can expect an *O* conclusion also to be often accepted; and vice versa.

'Caution' would lead to the acceptance of a particular conclusion when atmosphere favors the corresponding universal. Thus, with the premises *AA*, we find fully as many subjects accepting the *I* as the *A* conclusion; and similarly when atmosphere calls for an *E* conclusion we find large percents accepting the *O* conclusion as well. There is however excellent logical ground for such acceptances, once 'atmosphere' is recognized; for an *A* proposition logically implies its sub-

TABLES SHOWING THE EFFECT OF ATMOSPHERE ON ACCEPTANCE OF INVALID CONCLUSIONS

Entries indicate average percent accepting each type of false conclusion from given premises. Blanks mean that the corresponding types were not tried in the experiments.

TABLE I

($n = 81$)

Premises	Types of Invalid Conclusion			
	A	*E*	*I*	*O*
AA	**36**		33	11
AE	4	**45**	4	34
EE	3	**21**		13
AI	6		**62**	39
II			**51**	29
AO			33	**47**
EI		23	8	**44**
OO			18	**32**

TABLE II

($n = 90$)

Premises	Types of Invalid Conclusion			
	A	*E*	*I*	*O*
AA	**51**		66	27
AE	8	**51**	19	54
EA		**56**		59
EE		**41**		37
AI	18		**66**	
IA	21		**70**	
II			**68**	
AO		18	41	**71**
OA		21	43	**74**
EI		31		**58**
IE		32		**59**
IO				**60**
OI				**63**
EO				**53**
OE				**49**
OO				**54**

ordinate *I*, an *E* its subordinate *O*. Whenever a subject would have accepted an *A* conclusion, he should be willing to accept an *I* also. (He would not do so, of course, if and when he took *some* to mean *not all*.) Thus nearly all the

acceptances not referable to mere atmosphere can be explained by atmosphere plus a little logic, or by atmosphere weakened by caution, or by the ambiguity of *some*. The only remaining exception of any importance is the fairly large percent accepting conclusion *E* from premises *EI*, where atmosphere favors *O* and not *E*. A syllogism in which this occurred is quoted from Wilkins on page 452 above.

The order of the premises seems to have little effect on their atmosphere or on the acceptance of the conclusion. Table II affords a comparison of six pairs of premises presented in both orders. We observe that, with premises *AE*, 51 percent accepted a false *E* conclusion, as against 56 percent accepting this conclusion from premises in the reverse order, *EA*. The other comparisons show similar small differences.

An interesting corollary of the atmosphere hypothesis has to do with *valid conclusions*. Since the secondary hypotheses of atmosphere (page 453) correspond with certain formal rules of the syllogism, to the effect that if one premise is negative the conclusion must be negative, and if one premise is particular the conclusion must be particular, atmosphere should function to facilitate the drawing or acceptance of a valid conclusion. Therefore the valid conclusions should be less difficult than the invalid. In the Sells experiment with 300 syllogisms the average percent of errors for the 71 valid items was 16, while for the 229 invalid items it was 40. For 48 items presenting the Fallacy of Universal Conclusion from a Particular Premise, which proved to be the least difficult of the nine formal fallacies presented, the average percent of errors was 20.

A further bit of evidence for atmosphere can be drawn from the Wilkins data on acceptance of *valid* conclusions. Premises *AA* are capable of yielding valid conclusions in either *A* or *I*, and premises *AE* (or *EA*) can yield valid conclusions in *E* or *O*. Atmosphere should give the advantage to the *A* conclusions from *AA*, and to the *E* conclusions from *AE*, whether valid or invalid. The percents of acceptance are as follows:

Premises	AA			AE		
Conclusions	*A*	*I*	S.D.-diff.	*E*	*O*	S.D.-diff.
Valid	96	76	3.5	80	59	3.0
Invalid	36	33	3.2	45	34	4.9
Valid–Invalid	60	43		35	25	

Reading horizontally in this table we see that valid conclusions, as well as invalid, are more often accepted when favored by atmosphere than otherwise. Reading vertically we see that the valid conclusions are more often accepted than the invalid. The difference, Valid-Invalid (like the difference Right-Wrong in a True-False test) can be regarded as a measure of the amount of acceptance of the valid conclusions which is due to seeing the logical relationship. That is, while this group of college students showed a 96 percent acceptance of the valid *A* conclusions from *AA* premises, they also showed a 36 percent acceptance of invalid *A* conclusions from the same type of premises, because of atmosphere or other illogical factors, and therefore they can be credited only with a $96 - 36 = 60$ percent net score for genuine understanding of the relationships in *AAA* syllogisms. Their understanding was much less in *AEE* syllogisms.

The evidence for the atmosphere hypothesis falls short of what would be demanded for a complete demonstration. The gaps in the tables need to be filled in by further experiments now in progress. These gaps represent types of syllogisms which are usually considered too absurd to deserve a trial, and reference back to our first list of sample syllogisms will show how absurd many of them are. Still they should be tried and may prove to be more deceptive than they seem.

Summary

The problem is that of discovering psychological factors operating towards the acceptance of invalid conclusions in a syllogism test. Three such factors are suggested: the ambiguity of the word *some*, which is used in a distributive sense

in logic ('at least some') and very often in a partitive sense in ordinary speech ('only some'); 'caution' or wariness, favoring the acceptance of weak and guarded rather than of strong conclusions; and 'atmosphere,' the global impression or 'feel' of the premises, which is affirmative or negative, universal or particular. Examination of the data from two experiments indicates that nearly all the acceptances of invalid conclusions can possibly be explained by these three hypothetical factors.

(Manuscript received March 27, 1935)

REFERENCES

1. EIDENS, H., Experimentelle Untersuchungen über den Denkverlauf bei unmittelbaren Folgerungen, *Arch. f. d. ges. Psychol.*, 1929, **71**, 1–66.
2. STÖRRING, G., Experimentelle Untersuchungen über einfache Schlussprozesse, *Arch. f. d. ges. Psychol.*, 1908, **11**, 1–127.
3. STÖRRING, G., Allgemeine Bestimmungen über Denkprozesse und kausale Behandlung einfacher experimentell gewonnener Schlussprozesse, *Arch. f. d. ges. Psychol.*, 1925, **52**, 1–60.
4. STÖRRING, G., Psychologie der disjunktiven und hypothetischen Urteile und Schlüsse, *Arch. f. d. ges. Psychol.*, 1926, **54**, 23–84.
5. STÖRRING, G., Psychologie der zweiten und dritten Schlussfigur und allgemeine Gesetzmässigkeiten der Schlussprozesse, *Arch. f. d. ges. Psychol.*, 1926, **55**, 47–110.
6. STÖRRING, G., *Das urteilende und schliessende Denken in kausaler Behandlung*, Leipzig, Akademische Verlagsgesellschaft, 1926. Pp. xv + 232.
7. STÖRRING, W., Experimentelle Untersuchungen über einfache und komplexere Schlussprozesse, *Arch. f. d. ges. Psychol.*, 1925, **50**, 467–512.
8. WILKINS, M. C., The effect of changed material on ability to do formal syllogistic reasoning, *Arch. of Psychol.*, 1928, **16** (whole no. 102), pp. 83.
9. WILHELM, W., Über Mittel und Fehler beim Schliessen, *Arch. f. d. ges. Psychol.*, 1934, **92**, 175–192.

AN EXPERIMENTAL STUDY OF THE PAIRING OF CERTAIN AUDITORY AND VISUAL STIMULI [1]

BY JOHN T. COWLES

In spite of the accepted variabilities of effect of either auditory or visual stimuli and the supposed incompatibility, or at least complexity, consequent upon combining the two modes, equivalence in certain gross effects upon the observer seems fairly common in everyday life. The following experiment was directed to find what agreement there might be among individuals as to the correlation of music with graphic art, and to find any causes of such correlation.

Walker [2] has employed a combination of music and pictures in studying the relative musical development of children. Six simple melodies were played; the child was asked to characterize each verbally. Then, as each melody was played again, the child was asked to select from a group of six pictures the one which was most suitable. A prearranged correspondence was expected, such as a picture of marching soldiers with the 'March' melody, or the picture 'Mother at the Cradle' with the 'Lullaby' melody. Agreement with this (adult) correspondence was most frequent in the case of those children best able to describe the isolated melodies.

In the present experiment the subject listened to a musical selection played on a phonograph, and at the same time he viewed one by one a group of colored reproductions of paintings. Ten subjects observed a preliminary series of eight musical selections and eight pictures, in order to determine whether there was any significant agreement as to which picture aroused an affective mood most similar to each selection of music. Such agreement was found, even though the pictures in this series were simple landscapes; it ranged from

[1] From the Psychological Laboratory of Princeton University.

[2] Walker, E., *Das musikalische Erlebnis und seine Entwicklung.* Gottingen: Vandenhoeck, Ruprecht, 1927. Pp. 83, 91–96.

an agreement of seven subjects on each of two combinations to three subjects on each of three combinations. In this preliminary series and Series A the subject was instructed to choose the picture whose affective mood was most similar to that of the music being played, but in Series B the subject was asked merely to select the picture which "best fits the music;" he then recorded his basis of choice in a brief introspective account of what in the music and in the picture, or in neither, brought about such a correspondence. This last category, "in neither," gave an opportunity for the recording of such factors as extraneous associations, which might have affected the choice. After each choice, the picture selected was replaced in the group as a possible choice during succeeding musical presentations.

In the case of each choice, the subject was asked to record the degree of correspondence of picture and music, on a simple scale of three: well, fairly, poorly. This presumed that, while the picture which was chosen might fit only poorly, it would correspond at least to some degree. The subject was also asked to record his familiarity with the music and with the pictures.

The musical selections in all cases were instrumental music played on the phonograph. Music was selected which would probably be unfamiliar to most listeners. The selections differed in character, but within each one there were no more than slight or uniform variations in volume, tempo, or tone quality.

The sound source was placed behind the subject, who was seated at a table with a large screen of neutral color so placed as to exclude visual distraction. The experimenter controlled the phonograph in a different room; an electric pickup and radio amplifier with speaker made this possible. A buzzer and key allowed the subject to signal for change in musical selections. Needle scratch and distortion were minimal.

The pictures used were colored reproductions of paintings, averaging about 7½ in. by 9½ in., mounted on heavy white cards, uniformly 11 in. by 14 in. They were, for the most part, landscapes or scenes of simple content, with no limitation to a certain style. The cards were numbered in each series on the reverse side, and were viewed singly in regular and then reverse order for each musical selection. Although the subject controlled the time of exposure, there were but slight differences in total time for the various subjects. Fatigue was negligible, the experiment lasting for a half hour or less in each case.

Each series had a different group of subjects. They were undergraduates in Prince-

ton University, most of whom had some knowledge of psychological method but who were relatively unpracticed in introspection. As evidence of the unfamiliarity of the music throughout the experiment, in only 45 of a total of 512 cases was the music reported as more than slightly familiar. The number of pictures which had been seen before was negligible.

PHONOGRAPH RECORDS

I. Mars (Part I), from 'Planets' suite	Holtz	(Columbia)
II. Allegro	Fiocco	(Victor)
III. Siegfried Idyll (Part I)	Wagner	(Columbia)
IV. Marche Joyeuse	Chabrier	(Victor)
V. Tannhauser, Venusberg Music (Part I)	Wagner	"
VI. Symphony III, Marche Funebre (Portion)	Beethoven	"
VII. Die Meistersinger, Overture (Part I)	Wagner	"

Note: In Series B, for Record II was substituted 'Ballet des Sylphes' by Gounod (Victor).

PICTURES

1. Lake Geneva	Holder	(Artex Print)
2. A Northeaster	Homer	" "
3. Springtime and Love	Michetti	(Museum Print)
4. Die Sandgrube	Jaeckel	(Seeman Print)
5. Gewitterlandschaft	Caspar-Filser	" "
6. Dance of the Nymphs	Corot	(Museum Print)
7. Fleeing from the Flames	Schreyer	" "

TABLE I

SERIES *A*. CHOICES OF 26 SUBJECTS

Pictures	Musical Selections						
	I	II	III	IV	V	VI	VII
1	.0	3.8	80.8	.0	.0	11.5	11.5
2	57.7	.0	.0	3.8	46.2	7.7	26.9
3	3.8	19.2	.0	34.6	3.8	3.8	7.7
4	.0	.0	3.8	7.7	.0	46.2	11.5
5	7.7	3.8	15.4	15.4	23.1	23.1	30.8
6	.0	73.1	.0	15.4	.0	3.8	7.7
7	30.8	.0	.0	23.1	26.9	3.8	3.8

The figures in all tables are percentages.

TABLE II

REDUCED SERIES *A*. 17 SUBJECTS; MUSICAL SELECTIONS V, VII, PICTURES 5, 7 OMITTED

Pictures	Musical Selections				
	I	II	III	IV	VI
1	.0	.0	82.5	.0	29.4
2	82.5	.0	.0	17.7	17.7
3	.0	47.1	.0	17.7	11.8
4	11.8	.0	11.8	29.4	41.2
6	5.9	53.0	5.9	35.4	.0

TABLE III

SERIES *A*. 14 MUSICALLY TRAINED SUBJECTS

Pictures	Musical Selections						
	I	II	III	IV	V	VI	VII
1	.0	7.1	85.7	.0	.0	.0	7.1
2	57.2	.0	.0	7.1	50.0	7.1	35.8
3	.0	14.3	.0	42.8	.0	7.1	14.3
4	.0	.0	.0	14.3	.0	42.8	7.1
5	14.3	7.1	14.3	14.3	14.3	35.8	21.4
6	.0	71.4	.0	14.3	.0	7.1	7.1
7	28.6	.0	.0	7.1	35.8	.0	7.1

TABLE IV

SERIES *A*. 12 UNMUSICAL SUBJECTS

Pictures	Musical Selections						
	I	II	III	IV	V	VI	VII
1	.0	.0	75.0	.0	.0	25.0	16.7
2	58.3	.0	.0	.0	41.7	8.3	16.7
3	8.3	25.0	.0	25.0	8.3	.0	.0
4	.0	.0	8.3	.0	.0	50.0	16.7
5	.0	.0	16.7	16.7	33.3	8.3	41.7
6	.0	75.0	.0	16.7	.0	.0	8.3
7	33.3	.0	.0	41.7	16.7	8.3	.0

TABLE V

SERIES *B*. 33 SUBJECTS

Pictures (Same as Series *A*)	Musical Selections (As in Series *A*, except, II)				
	I	II	III	IV	VI
1	.0	24.2	63.6	.0	12.1
2	36.4	3.0	.0	57.5	3.0
3	6.1	12.1	6.1	6.1	.0
4	.0	.0	12.1	.0	48.5
5	12.1	6.1	12.1	12.1	36.4
6	9.1	54.5	6.1	.0	.0
7	36.4	.0	.0	24.2	.0

TABLE VI

SERIES *B*. 18 MUSICALLY TRAINED SUBJECTS

Pictures	Musical Selections				
	I	II	III	IV	VI
1	.0	33.3	55.5	.0	11.1
2	33.3	.0	.0	66.7	5.6
3	5.6	5.6	.0	5.6	.0
4	.0	.0	16.7	.0	61.1
5	16.7	.0	16.7	5.6	22.1
6	5.6	61.1	11.1	.0	.0
7	38.9	.0	.0	22.2	.0

TABLE VII

SERIES *B*. 15 UNMUSICAL SUBJECTS

Pictures	Musical Selections				
	I	II	III	IV	VI
1	.0	13.3	73.4	.0	13.3
2	40.0	6.7	.0	46.6	.0
3	6.7	20.0	13.3	6.7	.0
4	.0	.0	6.7	.0	33.3
5	6.7	13.3	6.7	20.0	53.4
6	13.3	46.6	.0	.0	.0
7	33.3	.0	.0	26.7	.0

DISCUSSION OF DATA

In Series *A* certain fairly high agreements appear, especially for combinations II–6, III–1, VI–4, I–2, and V–2. These last two combinations, in each of which about half of the subjects agreed, may indicate that the same factors in Picture 2 brought about a correspondence with both selections of music, or that both selections of music were similar in effect. It is perhaps fully as significant as the agreements upon certain combinations that 14 out of the 49 possible combinations of pictures and music were *not* selected. We might distinguish several degrees of exclusion of combinations. In the case of Music III, for example, we find a well defined mode with only two other pictures ever selected for combination. In the case of Music VI we find an agreement of nearly half the subjects upon Picture 4, but those who did not select that

picture chose every possible one. Here the mode was well defined but the scatter great. In the case of Music I, however, we find two modes with practically no scatter; and in the case of Music VII or Picture 5 we find scatter approaching complete disagreement, with no well defined mode. Hence there seems to be all degrees of agreement on the part of the subjects, from the cases of four musical selections in which a definite mode was established with but slight scatter to the cases of Music VI and VII or Picture 5, where every possible combination was selected.

The reduced Series *A* was simply a repetition of Series *A*, but with musical selections V and VII and Pictures 5 and 7 eliminated. This reduction in the number of musical selections and pictures was made in order to find out whether this might compel subjects, differing from each other but slightly in their interpretation of the mood effect of the music and pictures, to center their judgments on fewer combinations. The results are not conclusive, as will be seen from Table 2. A larger percentage of the subjects selected the combinations I–2 and III–1, and there appeared two, nearly equal and well defined modes for Music II, *i.e.* Pictures 3 and 6, where in the previous series there had been a single mode and considerable scatter. Different observers were used in order to avoid the influence of previous judgments. [Uncontrollable changes in attitude would have occurred.] Approximately the same proportion of pictures was never selected as in Series *A*, an even more significant fact than before, considering the smaller number of possible choices. In addition to this, five new combinations were selected in this series not chosen in Series *A*, *i.e.* I–4, I–5, III–5, V–1, and V–4. This choosing of new combinations only tended to offset a concentration of judgments. It is interesting to note that of these five new combinations, chosen 16 times, 4 choices were judged to correspond 'well,' which is approximately the same proportion of 'well' choices as in the other combinations of the reduced Series *A*, but quite different from the proportion of such choices in the full Series *A*. The greater freedom of choice in the full series might have been a factor here.

In Series *B* the instructions to the subject were altered: he was asked simply to 'select the picture which best fits the music,' and to report what seemed to lead to his choice. From the introspective reports of this series it was evident, despite individual differences of attitude, training, and perceptive processes, that when many subjects chose a certain combination their reasons for so doing were strikingly similar. And, of these more frequent combinations, wherever a picture was selected to correspond with a selection of music with prominent dynamic characteristics, such as marked rhythm, changes in loudness, peculiar tempo, it was the represented content of the picture which appeared most often in the various reports of the reasons for such correspondence. The adjectives applied to the music were most frequently those implying that particular motor activity which was represented in the picture. Furthermore, in the case of these selections of music with prominent dyanmic qualities, the pictures in practically all cases were interpreted in motor terms, similar or identical with those terms of rhythm, tempo, loudness changes, in which the music was described. Line and color of the picture were but rarely mentioned. There were also very few reports upon mood. For example, of the 24 choices of the combinations Music I—Picture 2, Music I—Picture 7, descriptions of dynamic qualities of the pictures occurred in all but 4 choices: in one of these the effect of color was noted, in the other three an affective reaction of mood. Similarly, of the 18 choices of Music II—Picture 6, the only subject not attributing correspondence to a *dance* rhythm in the music and a dance represented in the picture was one who reported a certain musical association.

On the other hand, in the case of those selections of music with less prominent, less changing dynamic characteristics, we find the pictures more often interpreted in terms of mood, as 'restfulness,' 'desolation,' or 'sadness,' and with more frequent mention of colors and lines.

With the exception of Music VII in Series *A*, it was found that those combinations in which 30 percent or more of the subjects concurred are such that, with music of conspicuous

motor qualities, only pictures with some content, capable of activity, are chosen; while with music of lesser dynamic qualities only those pictures are selected which have much less, more indefinite content, incapable of motor activity. With but very few exceptions, pictures with 'less content' were never chosen with 'dynamic' music; while pictures of 'content' were never chosen with 'less dynamic' music. It would appear that dynamic qualities, or their absence, are prominent as a deciding factor in the choice of a combination.

The reasons for the rarely chosen combinations were usually superficial associations suggested either by the music or the picture; several subjects admitted that there were definitely non-agreeing elements in the stimuli; some gave no definite reason or mentioned vaguely an aroused mood.

As to the frequency of choice of combinations in Series *B* as compared with Series *A* and the reduced Series *A*, we find a degree of correspondence comparable to that of the previous series. The modes were more definitely established as to pictures selected for musical selections IV and VI, with less scatter of judgments here even though there were more subjects than in any of the previous series. On the other hand, there was more scatter for Music I and III. Two or more pictures were never selected from the seven possible choices for each piece of music, which is slightly more than the proportion not selected in the previous series for the same selection. Any higher, general agreement in this series may be due to the fact that here, since an introspective report was required, the subjects had to employ more care in judging; and also because of the fact that in this series the subjects were not required to assume a particular attitude of affective response, to which they may have been unaccustomed and which they might have found artificial in the other series.

In Series *A* and Series *B* the subjects were classified into musical and unmusical groups, that is above or below the average in musical training and ability, so far as could be determined by interview. The tables III, IV, VI, and VII show that there is no significant difference for either series between the frequency of judgments of the two groups of

subjects as thus differentiated. The diversity of judgments is greater in the case of certain musical selection but less in the case of others, if we compare the two groups. This is true for both of these series. So few subjects were 'artistic' with respect to pictorial art that it was of no value to differentiate them as a separate group.

In the introspective reports of Series *B* we found practically no difference between these two types of subjects in the sort of reasons given. The associations noted by the musical more often included musical associations, such as 'folk songs,' 'Scheherazade Suite,' 'Grofe's Suite,' and so forth, than those associations recorded by the unmusical. Nevertheless, the two groups were quite undifferentiated as to ability to make an analysis of their choice of combinations.

Summary

1. Partial agreement was found in the choosing of certain pictures to correspond with certain selections of music. This agreement was essentially the same when the choice was to be based upon equivalence of mood as when the choice was simply one of correspondence and with an introspective report of reasons. Certain combinations were never chosen.

2. For a given combination the introspective reports agreed in a striking degree. Among the combinations most frequently chosen, pictures with represented content capable of motor activity were nearly always selected with the musical selections of prominent dynamic changes; and likewise, pictures of slight content were nearly always selected with music of relatively weak dynamic qualities. The introspective reports suggested this relationship. The less dynamic music and the pictures with less express content were most often said to correspond on a basis of mood or abstract elements. Formal elements of the pictures, however, were but rarely noted. Rhythm, tempo, and changes in loudness were most frequently noted in the music. There was no mention of hedonic response.

3. No substantial difference was found either in the choices or reports of musical and unmusical observers.

(Manuscript received June 12, 1934)

A PRELIMINARY STUDY OF BISECTION OF LINES [1]

BY LEONA CHIDESTER, M.A.

Menninger Clinic, Topeka, Kans.

The perception of the mid-point of a line is a relatively simple perceptual task, and one which has been the subject of many psychophysical researches. The early bisection experiments of Fechner, Volkmann, Chodin and others have been reviewed in some detail by Wolfe (**10**). A complete review of the literature will not be attempted here, but we shall indicate some of the general problems with which such studies have been concerned.

(1) Bühler (**1**) has pointed out that, in perceiving a line which is divided into two segments, one may have an immediate 'impression of proportion' which is surprisingly precise. He found the sensitivity to differences in proportion to be about twice as great as the sensitivity to differences in length between isolated lines.

In bisection experiments we have to do with a perceptual process which may perhaps be regarded as the simplest type of 'Gestalt' perception.

(2) Some experimenters have been especially concerned with the variable error of bisection as a measure of spatial sensitivity and with the dependence of the magnitude of this error on the length of the bisected line. There has been some controversy as to whether Weber's law applied here; most experimenters agree that it is approximately satisfied (*cf.* Wolfe (**10**)).

(3) Many investigations have paid particular attention to the constant errors of bisection. These constant errors are of some importance for the theory of visual space perception. Kundt (**6**), in 1863, reported some experiments on the bisection of horizontal lines in monocular vision. He found a

[1] The experiments described in this paper were performed in the Psychological Laboratory of the Ohio State University. The writer wishes to acknowledge the aid received from Professor F. N. Maxfield in planning and carrying out this work.

tendency to make the 'inner' half of a line (the half which is imaged on the temporal retina) shorter than the outer half, (the half which is imaged on the nasal retina). Hering (**5**) attributed this result to a non-uniformity in the distribution of retinal 'space-values.'

The work of later experimenters has shown that it is impossible to generalize Kundt's conclusions, which were drawn from experiments employing only one observer and one particular length of line. For example, Feilchenfeld (**2**) obtained the same result as Kundt with very long lines, while with shorter lines his constant error was practically zero. R. Fischer (**4**) reported that his constant errors were opposite in direction to Kundt's; when he attempted to bisect a line, he consistently made the outer half shorter than the inner half. Münsterberg (**7**), on the other hand, obtained still a different sort of result. His tendency was to overestimate a distance on the left, as compared with a distance on the right, no matter which eye was being used. But Wolfe (**10**) employing a large number of subjects, found that the prevailing tendency, at least in the case of long lines, was to overestimate right-hand distances. In Wolfe's experiments, free eye-movement was allowed, while the earlier experimenters employed steady fixation; also Wolfe's results applying to binocular rather than monocular vision. Ueno (**9**) has recently reported experiments resembling Wolfe's both in method and results.

The individual differences in length-perception which appear in bisection experiments seem to be significantly related to individual differences in depth-perception, and have been correlated with peculiarities in the form of the horopter (3).

(4) Systematic changes in the bisection-error with time have been studied by Wolfe (**10**) and Pearson (**8**). According to Wolfe, practice reduces the total error of bisection about one-quarter. Pearson found a tendency for the apparent mid-point to move towards the right, during a single experimental period. In addition to this 'sessional' change, there was a 'secular' shift towards the left throughout the whole series of experiments. If two sets of experiments followed

each other too closely, the sessional effect persisted into the next session.

Apparently Pearson did not employ a head-rest to keep his head at a constant distance from the lines. This may well have accounted for the sessional change which he found, if the distance of his head from the paper changed slightly but continually during the course of the experiments. His results will be discussed in connection with the data of the present study.

(5) Wolfe (**10**) has made a study of individual differences in bisection. He says: "Age produces less effect on the judgment of the middle of lines than might be expected. The total error of adults may be given as 1/80, and eighth-grade children are quite equal to adults. Fourth-grade children have an error of about 1/60. First-grade children show an error of about 1/40 for lines not longer than 200 mm. On the average there is practically no difference in the ability of men and women to mark the middle of horizontal lines. . . . The constant error in horizontal lines is largely an individual matter. In the longer lines, however, there is a decided tendency to overestimate the right half. In the shorter lines this tendency is less noticeable on the part of the men, while in case of the women it is reversed, that is, the left half is overestimated."

Method and Results

In the experiments which are now to be reported, special attention has been given to characteristic individual differences. At the same time, a comparison has been made between the results of binocular bisection experiments, and those of monocular experiments with each of the two eyes. In this connection, the factor of eye-dominance has been taken into account.

Six adults (D.A., S.B., E.D., M.P., X.Z. and L.C.) and three children (L.B., M.L., and M.B.) acted as subjects.

Five straight black lines of lengths 5, 10, 25, 37.5 and 50 cm were printed on large sheets of white bond paper. The paper was chosen so that the inking of the lines would be as uniform as possible. The sheets were 46.5 × 56.5 cm in size, in order that the end of no line would be close to the edge of the paper.

In every case the subject was seated facing the light to ensure symmetry of illumination. The sheets were placed in front of him, in such a manner that the line of sight was approximately perpendicular to the lines to be bisected. All lines, except the one being bisected, were covered, and the covering sheets were kept as far as possible from the line and parallel to it. Marks were made on the table, and after each line was bisected, the sheet was moved up so that the next line came opposite the marks. In this way, the inclination from the vertical was kept approximately constant.

A chin rest 50 cm long was employed, so that the distance of the eye from the line, and consequently the angle subtended by the line, remained constant. For the various lines this angle varied between 0.1 and 1.0 radian. No restriction was placed on eye-movements.

In each sitting, a set of 300 bisections was made. These sets were not interrupted for more than a few minutes' rest, except when absolutely necessary. Both eyes were first used together, then the right and then the left alone. The eye not in use was shuttered by a dark piece of cardboard fitted into a goggles frame.

The subject marked the apparent mid-point with a sharp pencil which he was nstructed to hold so that the point projected some distance beyond the hand, since the destruction of symmetry, when the hand was near, seemed to have some effect. All marks were made with the hand moving away from the subject in order that it might not cover the line at any time. The pencil was not held with the usual inclination to the right, but rather perpendicularly to the line which was to be bisected.

At least one complete set was secured from each subject, and for most of the subjects, three or four such sets. For each bisection, the left segment of the line was measured, as a percentage of the total length of the line. Values less than 50, therefore, indicate a deviation to the left, and those greater than 50 indicate a deviation to the right. The means of 20 observations of each of the five lines, for each of the three eye-conditions, were taken, and are given for each observer in the table. Figure 1 shows the means for each line and each eye-condition, as obtained for each subject, from all sets made by him.

The probable error of the mean was determined for each set of 20 observations. The means of these probable errors are given in the table, as well as means of them themselves for different groupings of conditions.

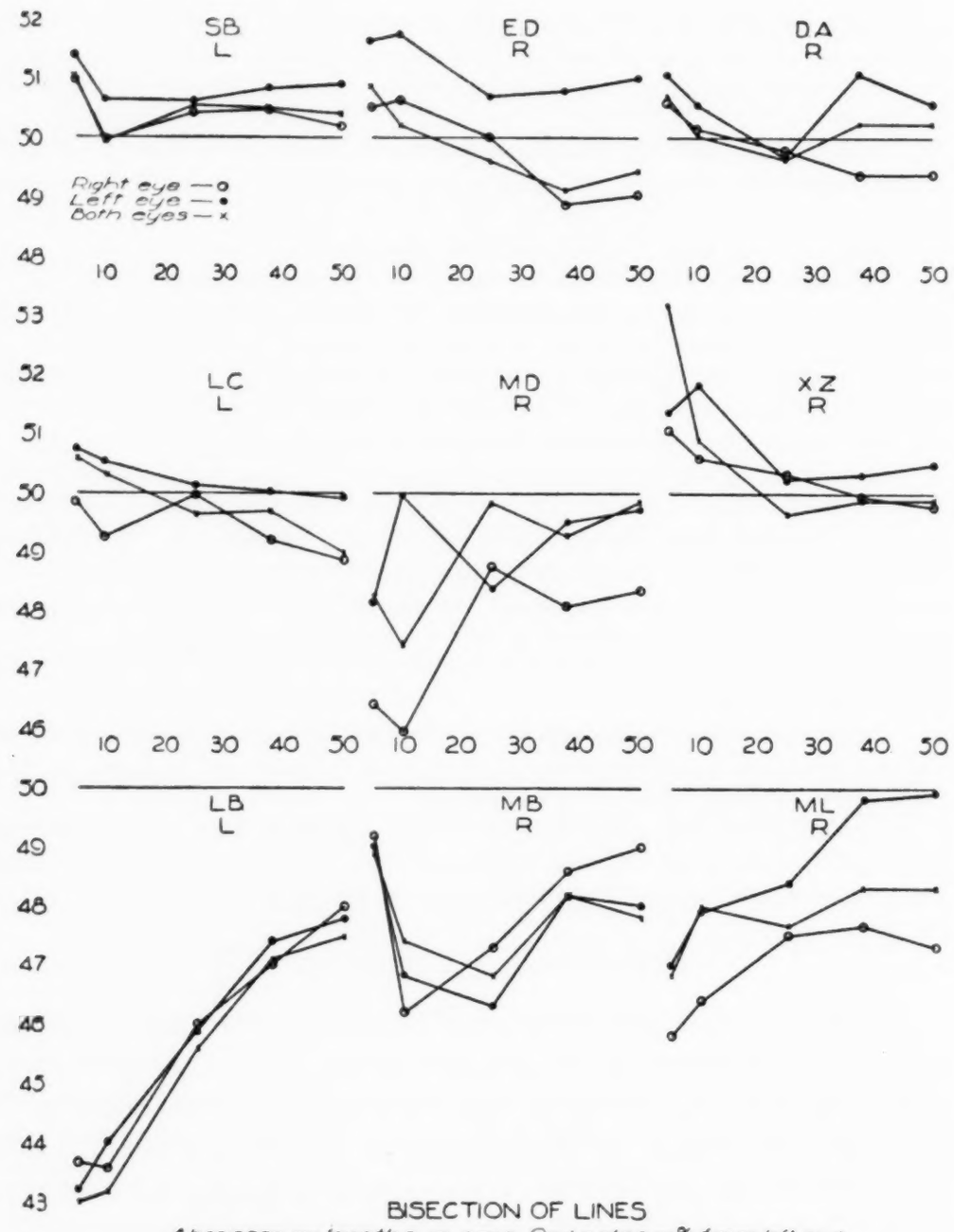
It generally has been assumed that both eyes together will bisect a line more accurately than either alone. To test this, the probable errors of each subject have been averaged for all five lines together and are shown as the bottom rows of each section of the table.

Results and Conclusions

1. There exist for lines of different lengths rather large systematic differences in the estimated points of bisection, made by any one subject. In some cases, these amount to 20 times as much as the probable errors involved. Any study based on one length of line only is seriously incomplete.

2. For any one line, the differences of estimation for different eye-conditions usually are large. This is especially true for the shortest and the longest lines, where the differences are sometimes 15 times the probable errors. A very striking exception to this rule is found in the case of the child, L. B.

3. In general, there is a strong resemblance between sets made by a given subject on different days, but there are also differences between these sets far greater than the probable errors determined for the set-means. These relationships



BISECTION OF LINES

Abscissae — lengths in cms. Ordinates — % from left end
R or L below initials indicates subject right or left eyed

verify the conclusions of Pearson (**8**), that there is a persistent bias of the subject through long intervals and that there are other biases which persist a few hours only.

4. The curves of means show that, in general, the results for the two eyes together are in no sense a mean of those for each singly. For S. B., the right eye and both eyes have closely similar curves (see Fig. 1, S. B.), while the curve for the left eye deviates strongly from the others. For most of

the other subjects, Fig. 1 shows that 'both' follows one of the eyes quite closely for short lines but may change to the other for long lines. In some cases, 'both' does not lie between the two monocular curves, but far to one side. This rule has exceptions in the case of young children.

5. Figure 3 shows that dispersions for both eyes are approximately the same as for each separately. This is surprising when one has listened to the remarks made by the subjects regarding the uncertainty felt by them when one eye is shuttered.

6. The facts stated in (4) and (5), although not proving this conclusion, indicate that bisection of lines is not at least for most subjects, primarily a binocular function. We have a rather closely analogous case to sighting a gun, which is a monocular function.

7. Ocular dominance does not seem to affect the dispersions. In four cases, the eye which is dominant, according to test with the Parson manoptoscope, has the smaller dispersion; in three cases the non-dominant eye has the smaller dispersion; and in two cases the dispersions are equal. This is at variance with the results obtained by Kundt (**6**), who says the errors of the right eye are twice those of the left. But it must be noted that he used only one length of line, 10 cm, and paid no attention to ocular dominance. Possibly Kundt's left eye was dominant. Moreover, as the data of this study will show, results that are true for one length of line or for one subject may well be contradicted by another set of measurements. In the case of E. D., there was disagreement between different criteria of dominance. According to the report of his oculist and according to our own tests with the manoptoscope, his right eye is dominant, yet he uses his left eye for sighting when at rifle practice. His binocular curves, especially for the longer lines, follow closely that of the right eye. In the case of four out of the other five adults the binocular curve follows the curve for the non-dominant eye, and follows it especially closely for the longer lines. Although we have too little evidence to draw certain conclusions, there is an indication that binocular bisection, in the case of adults, is usually controlled by the non-dominant eye.

In children, the slight amount of available evidence points to the reverse. This may indicate a phase of development of binocular vision. These two possibilities are merely mentioned here to show one direction which further investigation may take.

8. The results give a good deal of weight to an hypothesis that the adult uses the two eyes for quite different purposes. It may be indicated that he uses one habitually to grasp details within the field of vision and the second to place the image spatially. It is difficult otherwise to see why both eyes together should so commonly follow one of the eyes, instead of some sort of mean between the two separately. Especially is it difficult to see why the binocular estimates should follow the non-dominant eye with some exactness.

9. Figure 2 shows a feature that may be useful in mental testing. D. A. is a mature physicist, S. B. is a graduate student specializing in physics, and L. C. has had considerable experience in physics-laboratory work. For these subjects, the difference in dispersion between the short and long lines is far less than in the case of the other subjects, and the dispersions themselves are very low. For D. A., the curve of dispersions is almost a horizontal line. M. B., M. L., and L. B. are children with mental ages from 11 to 12 years and chronological ages from $9\frac{1}{2}$ to 11 years. In each of these three cases the dispersions are extremely large for the shortest lines, but decrease rapidly with length of line, until for the longest lines, they compare favorably with the dispersions of adults. (Note in Fig. 1 that the M. B., M. L., L. B., and X. Z. curves approximate more and more closely to the adult type of curve according to the order of their mental ages on the Stanford-Binet tests.) The analysis of the dispersions shows that Weber's law holds rather well, at any rate for the more highly trained adult, within the range of lengths used in this experiment. However, Weber's law breaks down entirely in the case of children.

10. The cases of M. P. and X. Z. are especially interesting. M. P. is a graduate student in modern languages, with an unusually high scholastic record. Her dispersions are inter-

mediate between those of a child and those of the subjects who have had experience in laboratory technique. X. Z. is a university senior specializing in social science, with an I.Q. (Stanford-Binet) between 90 and 95. She has great difficulty with college mathematics and has avoided exact sciences and laboratory work (knowing her own weakness). Her dispersions, and the shape of the curve derived from them, are almost identical with those of the three young children. It is impossible with the information at hand to decide between the hypothesis that the training received by laboratory experimentation decreases dispersion and the hypothesis that dispersion measures capability in that type of work It is hoped that it may be possible at a later date to have sets made by many students in physics-laboratory sections at the beginning and at the end of the term. Decrease in dispersion would establish the former hypothesis, while negative correlation between dispersions and success in the course would support the latter. It is quite important that this be done, since there is an indicated possibility for a test of a phase of intelligence which is not well considered in standard test. If a high correlation were found, it might be of value in selecting workers for certain mechanical trades.

11. With only two exceptions (both children) the apparent mid-points show a drift to the left between the 5 and 10 cm lines. For most subjects, there is a later drift to the right as we go to the longest lines. It would take observations of many more adults to establish this as a general rule. Attention must be called to L. B., aged 11 years, who starts with an extreme left bias and consistently moves farther toward the right, as longer lines are used. She not only deviates in this manner from adults, but also has almost identical curves for the three eye-conditions and extreme consistency in results from day to day. Tests of this remarkable child should be made each year, in order to study the way in which her estimations will approach those of adults.

12. Lack of ability to sustain attention is one of the characteristics commonly found with a low I.Q. Although the I.Q. was not determined for all of the adults, X. Z. is the

only one who would rate below 100 or 110. Her dispersions were not due in most cases to an unusual number of large errors, but to a few errors of extreme magnitude. For example, in set III, 5 cm, left eye, the mean of the set is 51.4 percent, but one of her values is 41.0 percent. The next largest error is about half that of this judgment. In watching her work, one could have no doubt of her intention to perform the task satisfactorily, yet her attention was so unstable that, although she systematically made the right segment shorter than the left, nevertheless in one case, she made that segment nearly 50 percent longer than the left one. This is all the more striking when one notes that 12 of the series of 20 judgments to which this case belongs had deviations from the mean of less than 2 percent. Sustained attention is a necessity for uniformly small dispersions in this problem, and it will be important in further work to determine how closely dispersions correlate with other criteria of the ability to sustain attention. It is interesting to note that in physical-laboratory work, undivided and continual attention is the first requisite for success.

(Manuscript received June 4, 1934)

References

1. Bühler, K., *Die Gestaltwahrnehmungen.* Stuttgart: Spemann, 1913.
2. Feilchenfeld, H., Ueber die Grössenschätzung im Sehfeld, *Graefes Arch. f. Ophth.*, 1901, **53**, 401.
3. Fischer, F. P., Ueber Asymmetrien des Gesichtsinnes, speziell des Raumsinnes beider Augen, *Pflügers Arch. f. d. ges. Physiol.*, 1924, **204**, 203.
4. Fischer, R., Grössenschätzungen im Sehfelde, *Graefes Arch. f. Ophth.*, 1891, **37** (1), 97.
5. Hering, E., *Beiträge zur Physiologie.* Leipzig: Engelmann, 1861–64.
6. Kundt, A., Untersuchungen über Augenmass und optische Täuschungen, *Poggendorffs Ann.*, 1863, **120**, 118.
7. Munsterberg, H., Augenmass, *Beitrage zur experimentellen Psychologie*, Heft 2, 125. Freiburg: Mohr, 1889.
8. Pearson, E. S., On the further investigation of the personal equation, *Biometrika*, 1922, **22**.
9. Ueno, J., Experiments in perceptual judgment of space bisection, *Jap. J. Psychol.*, 1926, **1**, 453.
10. Wolfe, H. K., On the estimation of the middle of lines, *Amer. J. Psychol.*, 1923, **34**, 313.

General Summary of Observations

Data given are means for each set of 20 observations of the percent of the whole line included in the left hand estimate. Initials refer to observer. The probable errors given are the means of the probable errors of the individual sets of 20 observations.

	Observer								
	S. B. Dominant eye left			L. C. Dominant eye left			D. A. Dominant eye right		
	Both	Right	Left	Both	Right	Left	B th	Right	Left
5 cm									
Set I	51.16	50.88	51.82	50.94	49.40	51.02	50.86	50.81	51.42
II	51.12	50.96	51.42	49.98	50.79	49.14	50.80	50.53	51.01
III	50.93	51.04	51.28	50.62	49.40	52.15	50.48	50.46	50.75
IV	50.91	51.08	51.14	50.80	49.85	50.65			
Mean	51.03	50.99	51.42	50.58	49.86	50.74	50.71	50.60	51.06
P. E.	0.145	0.139	0.152	00.144	0.172	0.166	0.075	0.09	60.082
10 cm									
Set I	50.28	50.25	50.82	50.02	48.96	49.72	49.88	50.03	50.50
II	49.75	49.94	50.58	51.54	50.55	51.52	50.20	50.23	50.68
III	49.80	49.96	50.66	50.18	48.83	51.03	50.09	50.22	50.50
IV	49.90	49.78	50.57	49.55	48.62	49.78			
Mean	49.93	49.98	50.66	50.32	49.24	50.53	50.06	50.16	50.56
P. E.	0.092	0.117	0.129	0.131	0.141	0.147	0.070	0.084	0.084
25 cm									
Set I	50.76	50.64	50.78	50.06	50.15	50.76	49.64	49.80	49.88
II	50.51	50.25	50.61	49.64	50.18	50.08	49.72	49.68	49.62
III	50.52	50.59	50.81	49.98	50.15	50.02	49.62	49.96	46.64
IV	50.46	50.28	50.36	48.89	49.35	49.64			
Mean	50.56	50.44	50.64	49.64	49.96	50.12	49.66	49.81	49.71
P. E.	0.106	0.111	0.110	0.116	0.148	0.146	0.082	0.093	0.088
37.5 cm									
Set I	50.42	50.94	51.07	50.16	49.53	50.79	50.59	49.20	50.95
II	50.48	50.49	50.71	49.82	49.46	49.64	50.12	49.36	50.84
III	50.50	50.70	51.30	49.28	49.12	49.84	50.03	49.64	51.46
IV	50.64	49.83	50.30	49.46	48.68	49.90			
Mean	50.51	50.49	50.84	49.68	49.20	50.04	50.25	49.40	51.08
P. E.	0.147	0.140	0.150	0.136	0.128	0.144	0.095	0.104	0.111
50 cm									
Set I	50.07	49.87	50.82	49.32	49.18	51.34	50.39	49.40	51.07
II	50.33	50.50	50.84	48.48	48.80	48.60	49.91	49.14	50.27
III	50.70	50.32	51.30	48.73	48.66	49.55	50.44	49.68	50.46
IV	50.36	50.19	50.62	49.32	48.77	50.88			
Mean	50.36	50.22	50.90	48.96	48.85	49.91	50.25	49.41	50.60
P. E.	0.119	0.144	0.113	0.138	0.155	0.145	0.109	0.091	0.076

GENERAL SUMMARY OF OBSERVATIONS—*Continued*

	Observer								
	E. D. Dominant eye right			M. P. Dominant eye right			X. Z. Dominant eye right		
	Both	Right	Left	Both	Right	Left	Both	Right	Left
5 cm									
Set I	50.76	50.38	51.86	48.25	46.40	48.15	52.15	52.95	52.65
II	50.91	50.86	51.52				55.50	49.50	50.00
III	50.94	50.38	51.56				51.85	50.65	51.40
Mean	50.87	50.57	51.65				53.17	51.03	51.35
P. E.	0.175	0.182	0.178	0.303	0.207	0.334	0.283	0.319	0.477
10 cm									
Set I	50.40	50.92	51.80	47.40	45.95	49.98	50.30	50.90	51.05
II	50.10	50.18	51.48				51.85	49.42	51.48
III	50.16	50.85	52.06				50.50	51.38	52.92
Mean	50.22	50.65	51.78				50.88	50.57	51.81
P. E.	0.193	0.206	0.204	0.268	0.300	0.295	0.324	0.293	0.333
25 cm									
Set I	49.36	49.98	51.09	49.85	48.76	48.39	49.96	50.78	50.70
II	49.56	49.85	50.34				50.11	49.55	49.80
III	49.98	50.17	50.64				48.84	50.59	50.18
Mean	49.64	50.00	50.70				49.64	50.31	50.23
P. E.	0.134	0.156	0.141	0.181	0.220	0.204	0.233	0.155	0.206
37.5 cm									
Set I	48.37	48.91	51.26	49.26	48.06	49.50	49.70	50.74	50.40
II	49.35	48.87	50.16				49.64	49.18	48.85
III	49.48	48.81	50.94				50.28	49.90	51.59
Mean	49.13	48.86	50.79				49.87	49.94	50.31
P. E.	0.165	0.169	0.196	0.202	0.155	0.197	0.194	0.130	0.158
50 cm									
Set I	49.48	49.06	50.64	49.85	48.34	49.72	49.50	50.66	50.64
II	49.24	48.98	50.96				49.04	48.68	49.16
III	49.70	49.10	51.41				51.06	49.96	51.68
Mean	49.47	49.05	51.00				49.87	49.77	50.49
P. E.	0.169	0.176	0.172	0.153	0.141	0.155	0.206	0.120	0.163

GENERAL SUMMARY OF OBSERVATIONS—*Continued*

	Observer								
	L. B. Dominant eye left			M. B. Dominant eye right			M. L. Dominant eye right		
	Both	Right	Left	Both	Right	Left	Both	Right	Left
5 cm									
Set I	43.4	44.7	43.3	50.4	51.6	50.0	46.8	45.8	47.0
II	41.6	43.0	43.2	47.4	46.8	48.0			
III	43.9	43.4	43.0						
Mean	43.0	43.7	43.2	48.9	49.2	49.0			
P. E.	0.377	0.411	0.342	0.502	0.716	0.542	0.624	0.583	0.663
10 cm									
Set I	43.6	43.7	44.3	48.8	46.9	46.5	48.0	46.4	47.9
II	42.6	42.9	43.4	45.9	45.6	47.0			
III	43.5	44.3	44.9						
Mean	43.2	43.6	44.0	47.4	46.2	46.8			
P. E.	0.317	0.297	0.290	0.368	0.373	0.496	0.413	0.212	0.416
25 cm									
Ser I	46.7	47.0	47.4	47.3	49.4	47.0	47.7	47.5	48.4
II	44.4	44.1	45.2	46.3	45.2	45.6			
III	45.6	47.0	45.1						
Mean	45.6	46.0	45.9	46.8	47.3	46.3			
P. E.	0.174	0.197	0.211	0.307	0.296	0.296	0.238	0.309	0.258
37.5 cm									
Set I	48.5	47.7	48.8	48.5	50.1	49.3	48.3	47.7	49.8
II	45.8	45.6	46.6	48.0	47.1	47.1			
III	47.0	47.6	46.8						
Mean	47.1	47.0	47.4	48.2	48.6	48.2			
P. E.	0.189	0.194	0.197	0.249	0.280	0.277	0.245	0.214	0.246
50 cm									
Set I	47.9	48.2	48.0	47.5	50.5	49.3	48.3	47.3	49.9
II	47.1	47.2	47.5	48.2	47.6	48.3			
III	47.6	48.7	48.0						
Mean	47.5	48.0	47.8	47.8	49.0	48.0			
P. E.	0.150	0.186	0.158	0.256	0.244	0.252	0.174	0.201	0.146

THE FROIS–WITTMANN PICTURES OF FACIAL EXPRESSION

BY WILBUR S. HULIN AND DANIEL KATZ
Princeton University

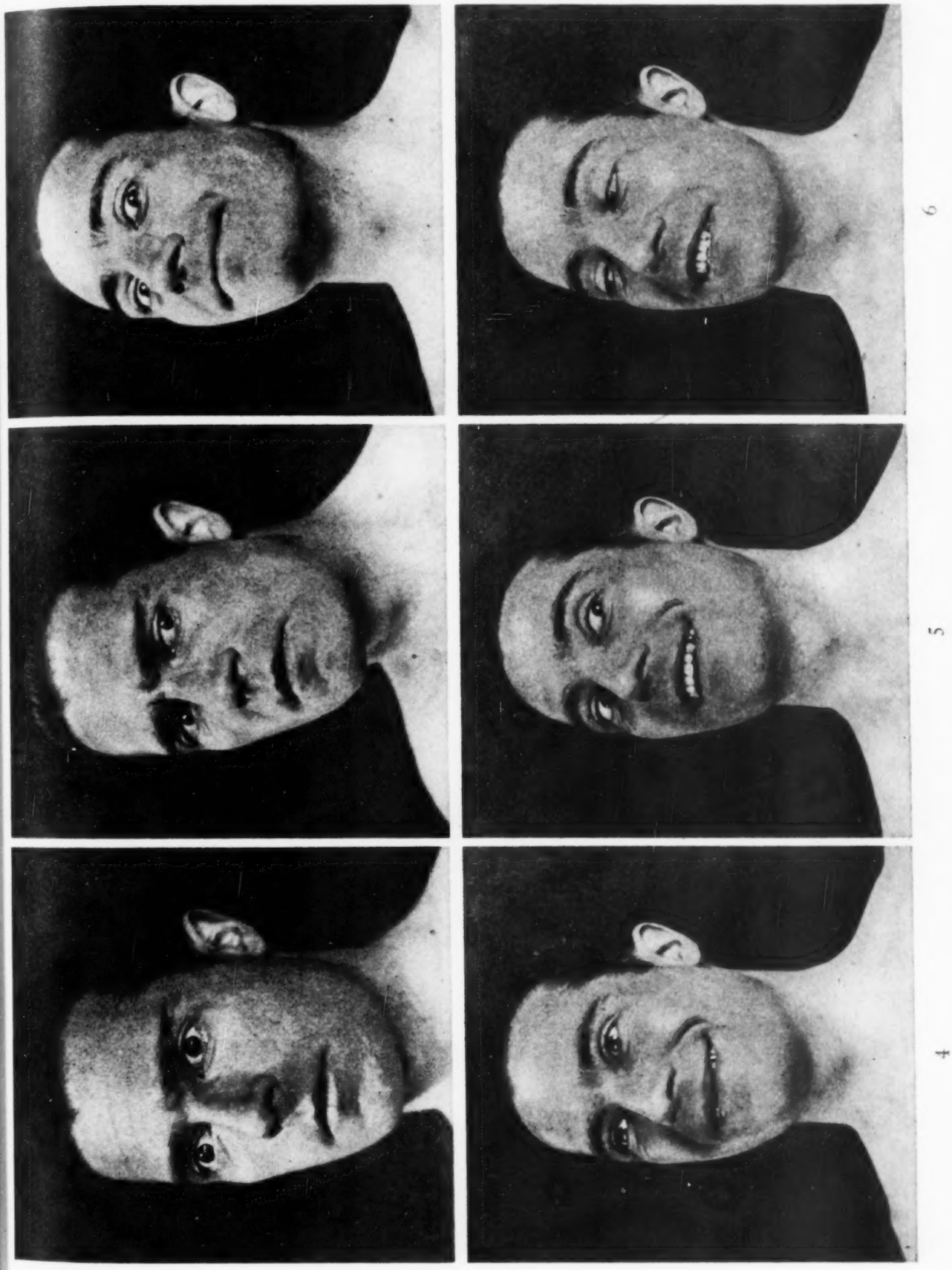
I. *Description of the Pictures.*—J. Frois-Wittmann, in his study of the judgments of facial expression,[1] made use of a new and extensive series of photographs. Frois-Wittmann posed for the photographs and designed the series to meet certain standards of uniformity in the pictures.[2] In all of the portraits the head is held in a three-quarter position. The head is tilted only when necessary for certain expressions. The face is clean-shaven, permitting the greatest possible evidence of the facial muscles. There is no indication of clothing. There are no gestures by other bodily members. The facial expressions assumed by the model were those most natural to him or were imitations of those shown in other series of facial expressions which had already given significant results. The expressions were always assumed voluntarily and practised in a mirror before photographing. The portraits were intended to represent a variety of possible and likely expressions, but were not intended to serve as predetermined standards of expression of any particular emotional or subjective states; the pictures were merely stimuli to be characterized by the judgments of the observers.

On the following pages are presented seventy-two of Frois-Wittmann's portraits. In the arrangement of the pictures no significant relationship is intended between adjacent pictures or between the pictures on a page. Also, the numbering of the pictures is arbitrary.

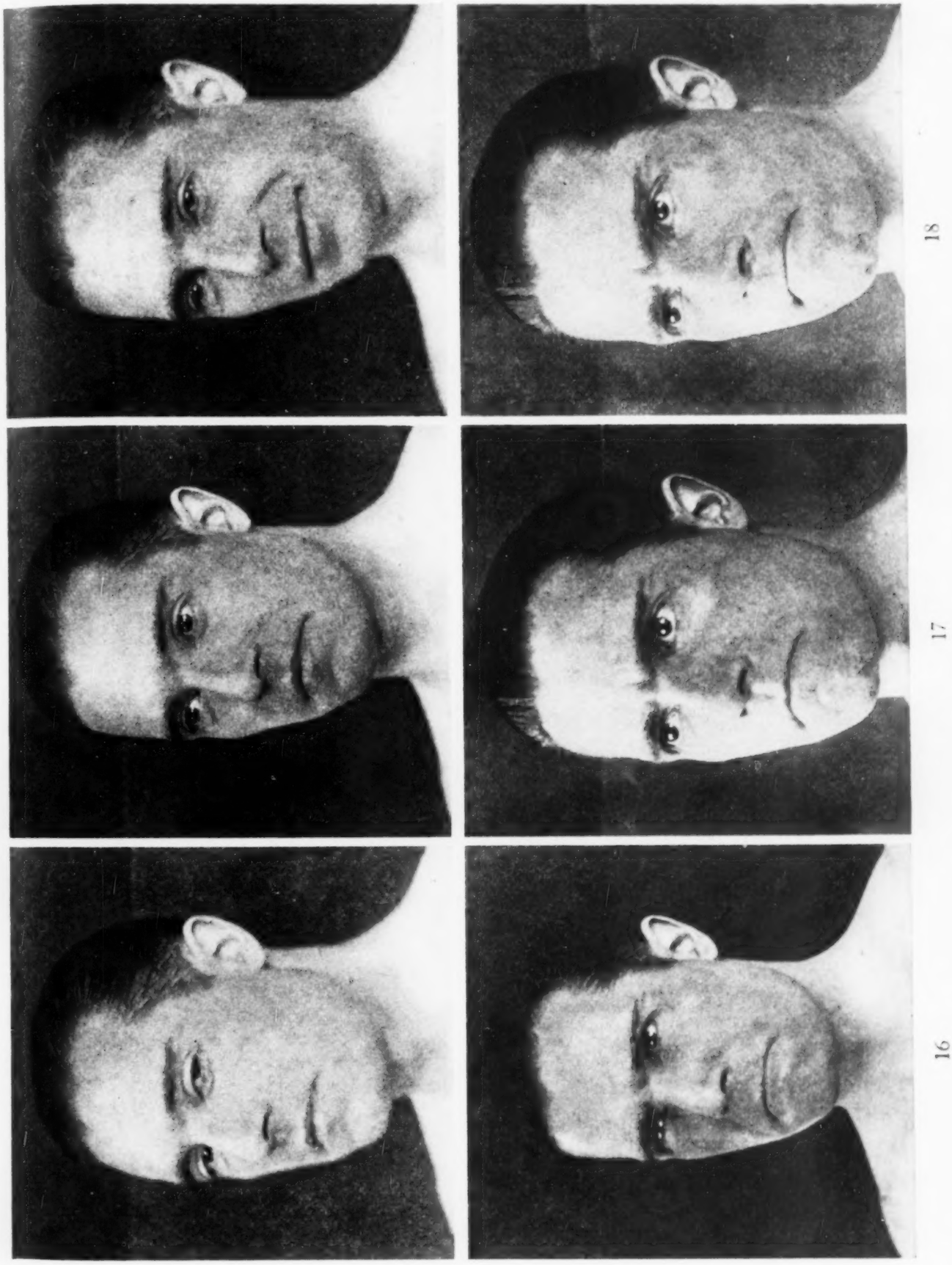
In Frois-Wittmann's study of the judgment of facial expression only forty-one of the photographs were retained for the main experiments. In addition to these, Frois-Wittmann employed pen and ink sketches of the photographs in

[1] J. Frois-Wittmann, The judgment of facial expression, this JOURNAL, 1930, **13**, 113–151.

[2] The photographs were taken by Professor Harold Schlosberg.


6
5
4

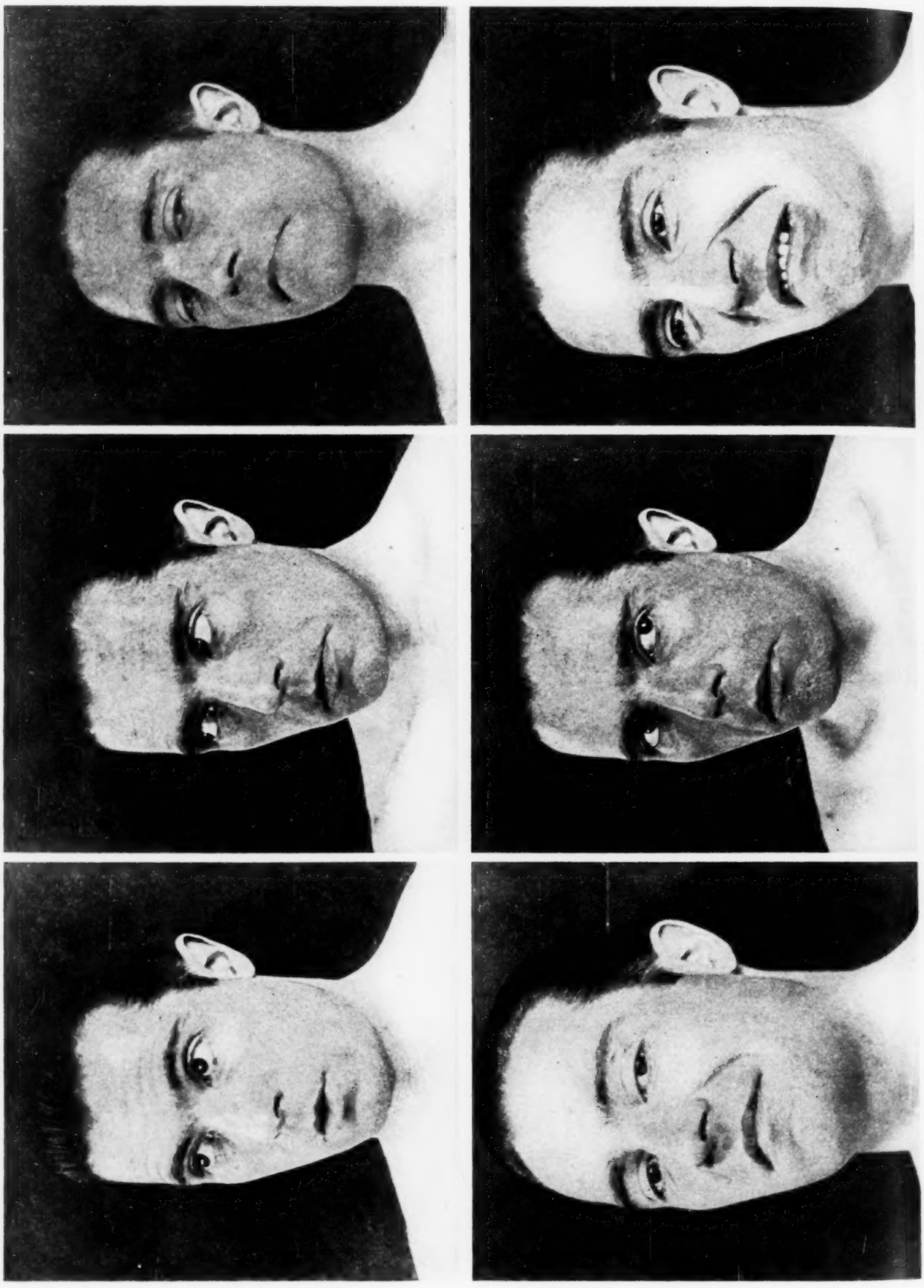
9
8
7



18

17

16


21
20
19

30

29

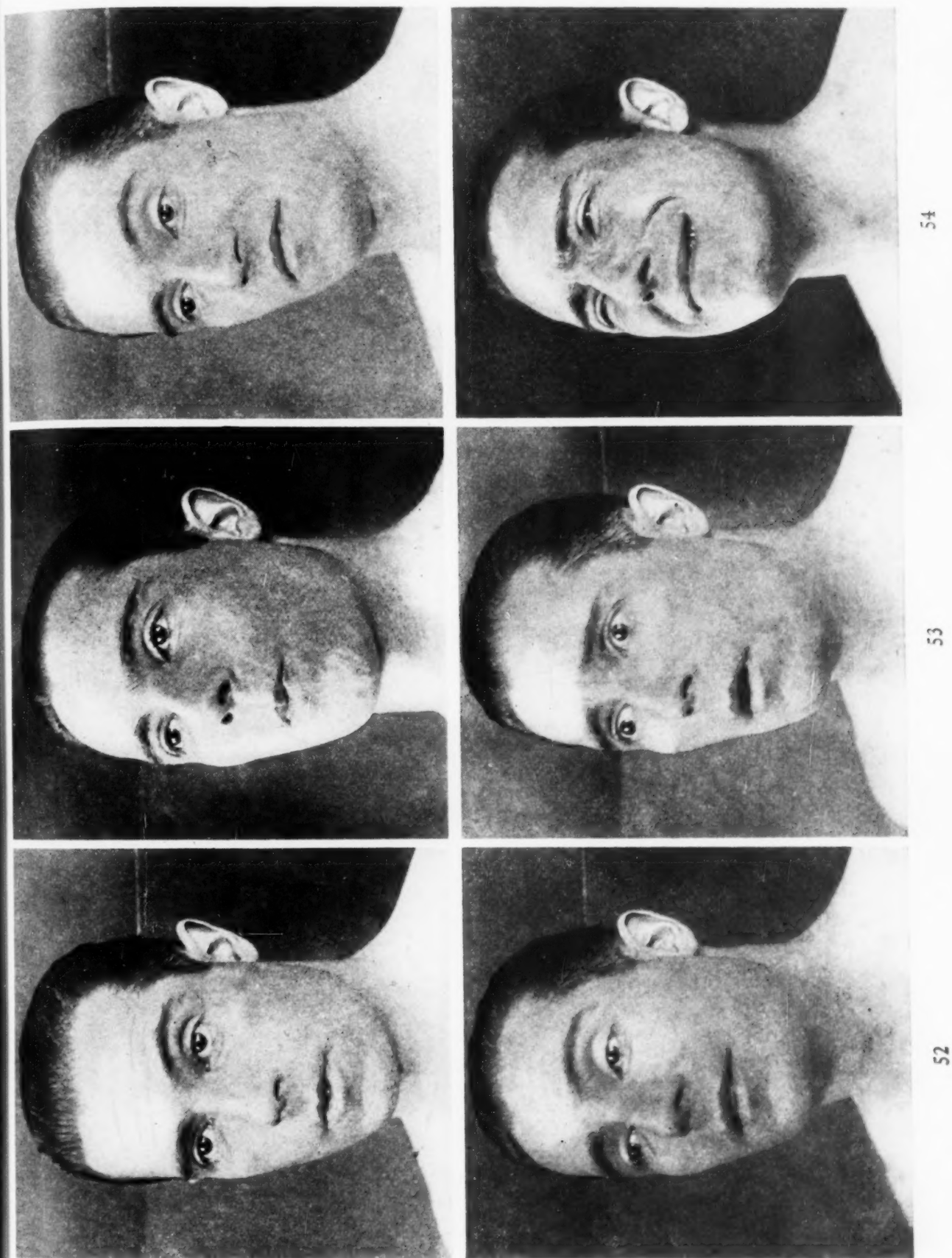
28

42

41

40

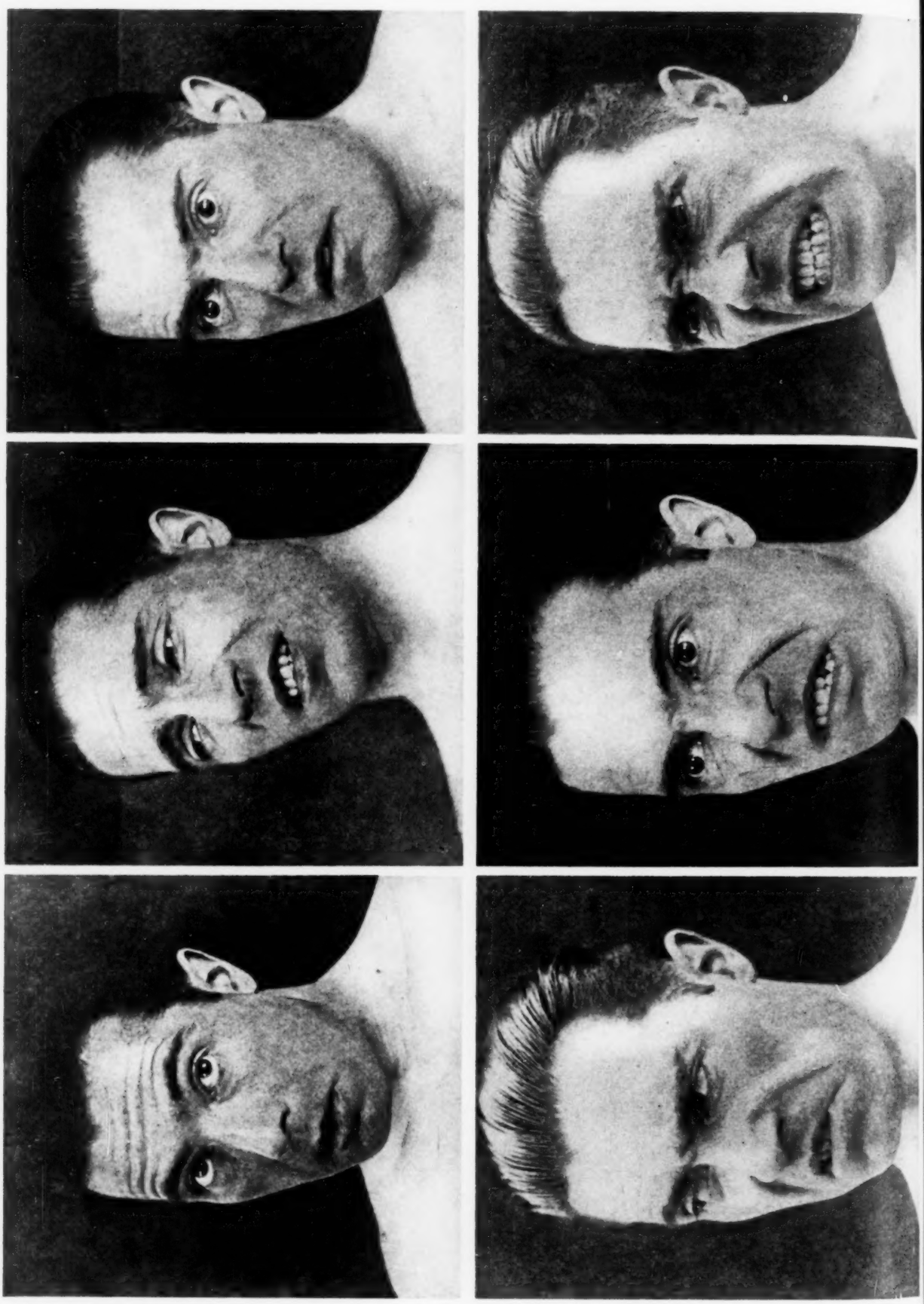
45
44
43



54

53

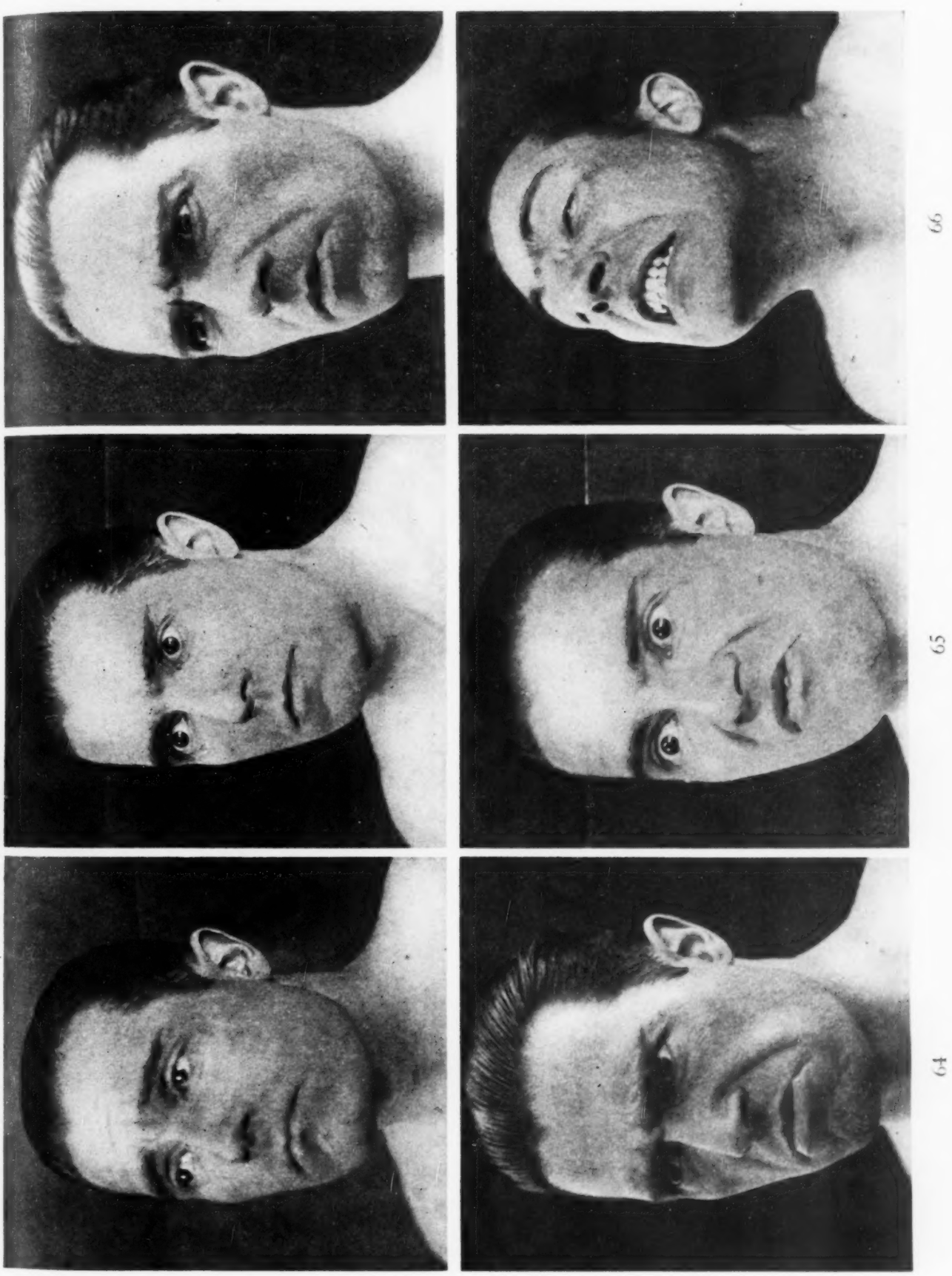
52



57

56

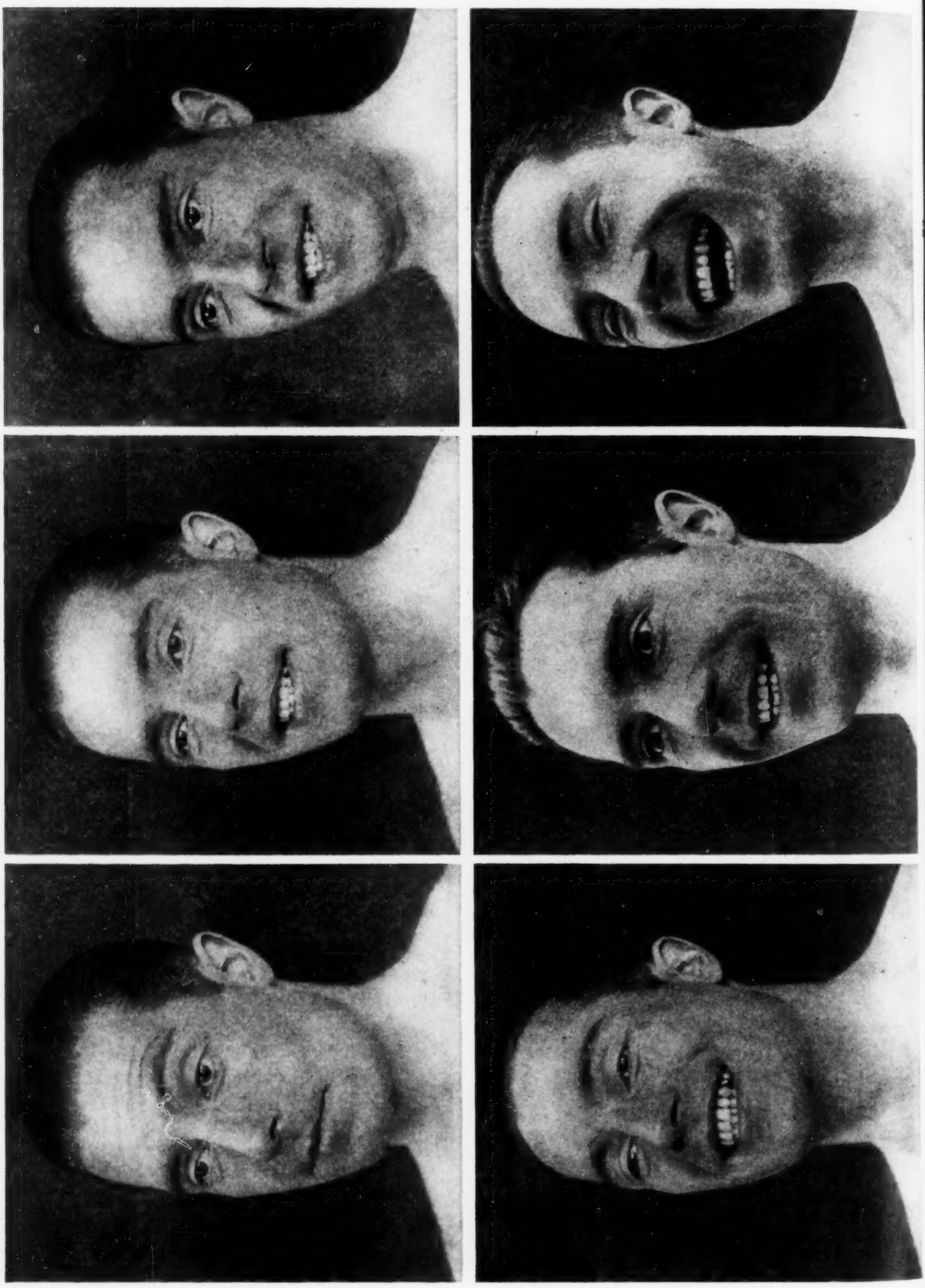
55



66

65

64



69

68

67

some of which the individual features were combined into composite pictures. Only ten photographs and three sketches were published by Frois-Wittmann. The remainder of the pictures have never appeared. In consideration of the range and uniformity of the photographs it seems worth while to publish a larger number of them.

II. *A Study of Non-verbal Judgments of Facial Expression.* —In Frois-Wittmann's study of the judgments of facial expression the observers reported their judgments of the pictures in terms of descriptive names, such as 'anger,' 'dismay,' etc. The pictures to which the same name was applied most frequently were then grouped together. Also, relationships between these groups were shown through the presence in different groups of the same descriptive names which appeared less frequently than the modal description. The value of this method of grouping pictures according to descriptive names has been questioned [3] on the grounds that the names are used differently by different observers and therefore the names themselves represent variable items upon which no reliable grouping of the pictures should be based. In view of the possible differences in meaning attached to the descriptive names the alternate situation arises for consideration, namely, the direct matching of pictures without the mediation of language.

In the present study the relationships judged to exist between the seventy-two pictures shown above were investigated by means of a method of sorting the pictures into groups, each group representing a distinct type of expression. The pictures were cut apart from each other so that they could be handled separately. There were twenty-two observers. The following instructions were given:

> Arrange the 72 photographs into groups according to the number of distinct types of facial expression which you judge to be represented. Place together in one group those photographs which approximate the same facial expression. Make as many separate groupings as you think proper. Your groupings do not have to contain the same number of photographs.

The results derived from the sorting of the pictures by

[3] *Cf.* C. C. Brigham, *A Study of Error*, New York, 1932, p. 50.

the 22 observers revealed both a wide scatter in the judgments and a few cases of high agreement. Table I shows the range

TABLE I

RANGE OF AGREEMENT ON SIMILARITY BETWEEN INDIVIDUAL PICTURES (*i.e.*, THE SORTING OF ANY TWO PICTURES INTO THE SAME GROUP)

Percent of Observers Making Same Judgment	Number of Times Same Judgment Made	Percent of Observers Making Same Judgment	Number of Times Same Judgment Made
100.0%	1	50.0%	19
95.5	2	45.5	20
90.9	3	40.9	29
85.5	7	36.7	27
81.8	4	31.8	32
77.3	5	27.3	46
72.7	8	22.7	49
68.2	6	18.2	84
63.6	6	13.6	100
59.1	11	9.0	195
54.5	16	4.5	382

of agreement among the observers on the similarity between any two pictures. This agreement is based upon the number of individuals who sorted any two pictures into the same group. All observers agreed upon two pictures as belonging in the same group in only one case, namely, pictures No. 27 and No. 29. In 382 cases (out of a total of 1,049) only one observer judged two pictures as similar.

The cases in which judgments of similarity between pictures were agreed upon by 50 percent or more of the observers are shown in Chart I. The top horizontal line of numbers and the left-hand vertical column are the numbers of the pictures involved in the agreements of 50 percent or more of the observers. (The numbers of all other pictures are omitted.) The figures shown within the chart indicate the percent of observers agreeing upon the similarity of the two pictures. For example, 68.2 percent of the observers agreed upon the similarity of pictures 3 and 5, *i.e.*, 68.2 percent of the observers grouped pictures 3 and 5 together.

The foregoing data are presented as an initial step in the establishment of standards for the Frois-Wittmann pictures. These data afford no direct basis for comparison with Frois-

CHART I

SHOWING THE CASES IN WHICH JUDGMENTS OF SIMILARITY WERE AGREED UPON BY 50 PERCENT OR MORE OF THE OBSERVERS

Picture	5	6	11	15	17	18	21	22	23	24	27	28	29	30	38	39	41	42	44	45	46	47	48	51	53	54	56	57	59	60	64	66	67	68	69	70	71	72
3	68.2	59.1																																				
4																																		50.0		50.0	50.0	
5		68.2																																				
6			54.5																																			
9																	59.1			59.1		54.5																
12								59.1		54.5																												
13				50.0			63.6	54.5																														
15																								54.5									54.5					
16					85.5	85.5									50.0	63.6															50.0							
17						90.9									54.5	72.7															50.0							
18															50.0	68.2																						
20									59.1																													
21								72.7																														
26											85.5	81.8	85.5						59.1						77.3													
27												95.5	100.0	50.0					50.0						68.2													
28													95.5	50.0					54.5						77.3													
29														50.0					50.0						72.7													
30																									54.5													
32																																		68.2	59.1	72.7	63.6	85.5
33																		77.3								81.8						59.1					
35																							63.6															
36																													81.8	90.9	50.0							
38																54.5																						
39																															50.0							
41																			50.0																			
42																						54.5				77.3						59.1					
44																					54.5				54.5													
45																						54.5					50.0					54.5					
46																										59.1		54.5				63.6					
47																																	50.0					
51																																72.7						
54																														72.7								
59																															59.1							
60																																		72.7	85.5	81.8	72.7
68																																			63.6	68.2	50.0
69																																				90.0	85.5
70																																					77.3
71																																					

Wittmann's results since they are derived from a different method of judgment and a different statistical treatment. The present article introduces the Frois-Wittmann pictures as an additional source of material in the study of facial expression.[4]

(Manuscript received May 23, 1934)

[4] For a summary of studies of facial expression see A. Jenness, The recognition of facial expressions of emotion, *Psychol. Bull.*, 1932, **29**, 324-350.

THE AFTER-EFFECT OF THE PERCEPTION OF CURVED LINES

BY J. F. BALES AND G. L. FOLLANSBEE[1]

I. Statement of the Problem

The after-effects of the perception of curved lines were first studied by J. J. Gibson.[2] Gibson found that after inspecting a black curved line on a white background for a certain length of time, a subsequently seen straight line would appear curved in the opposite direction. This phenomenon was observed both when the stimulus line was fixated and when the eyes moved up and down the stimulus line during the inspection period. During the inspection period the stimulus curve seemed to straighten somewhat and the amount of the after-curvature corresponded very closely to the amount that the stimulus line seemed to straighten. The after-effect was observed following inspection periods of one to two minutes although five to ten minutes gave the best results. The amount of curvature of the stimulus line did not greatly alter the after-effect. The phenomenon was found to be limited to the stimulated area of the visual field; however, transfer to the corresponding area of the other eye occurred.

Gibson carried out several experiments in an attempt to throw light on the underlying basis of this phenomenon. He concluded that the explanation of the after-effect is not to be found in secondary factors such as kinæsthesis, eye movement, unconscious inference, etc., but must lie in the nature of perception itself.

It is the purpose of this paper to verify the phenomenon of negative after-curvature as originally reported by Gibson and to study the duration of the after-effect under varying conditions.

[1] From the Psychological Laboratory of Princeton University.

[2] J. J. Gibson, Adaptation, after-effect and contrast in the perception of curved lines. J. Exper. Psychol., 1933, 16, 1–31.

II. Procedure, Apparatus, and Subjects

The subject was seated in a chair with his head held in a constant position by means of a headrest. Directly in front of him at a distance of 40 cm was a black curved line. He was asked to inspect this stimulus line, allowing his eyes to move freely up and down it. After the inspection period the stimulus was removed and a straight line was presented as a test. The subject was asked whether the test line appeared straight or curved. If he stated that it appeared curved, the test line was bent until the subject declared that it appeared straight to him. The extent to which the test line was bent gave a measurement of the amount of apparent curvature of the test line.

The stimulus line was a black curved line 3 mm wide drawn on a white cardboard (35 cm × 72 cm). It was 30 cm long with a displacement of 40 mm at the center convex to the left. The test line was a straight black flexible rod 32 cm long, held at the ends by two small white blocks of wood 30 cm apart. These blocks were firmly attached to a white cardboard background on an ordinary drawing board. A thin white wire attached to the center of the flexible rod was passed through a hole in the board so that but two inches of the wire could be seen. By pulling this wire the test rod could be bent.

The amount of after-curvature was measured on a millimeter scale on the back of the drawing board. Against this scale a marker attached to the wire indicated the amount of displacement of the center of the test line.

The stimuli and all accessory apparatus were entirely screened in by white curtains so that distracting influences were minimized. The screen also allowed the experimenter to measure the results out of sight of the subject.

A total of 43 subjects were used in the entire series of experiments. 39 were departmental students in Psychology; 26 had a slight knowledge of after-effects in color vision and elsewhere; and 6 had a slight previous knowledge of the particular experiment. No significant difference relating to knowledge of after-effects was found among the subjects.

III. The Experiments

Series I— Negative After-Effect of Curvature

Eleven subjects inspected the stimulus line for a period of 10 min. The after-effects as measured by the apparent displacement of the test line are given in Table I. This Table

TABLE I

The Amount of Displacement (in Millimeters) of the Center of a Straight Line Following the Inspection of a Curved Line for 10 Min.

Subject	Test I	Test II
1	1.75 mm	2 mm
2	2.5	2.5
3	2	2
4	0	0
5	2.25	2.5
6	2	2.5
7	2.25	2.25
8	0	0
9	2.25	2
10	2.25	2
11	1.75	1.5
Average	1.727 mm	1.75 mm

also gives the results of a second test made one week later. The average apparent displacement was 1.73 mm on the first test and 1.75 mm on the second. This closely approximates the average displacement of 1.67 mm found by Gibson.

The experiment was repeated with 11 new subjects, the inspection period being reduced to 5 min. A second test of the same subjects was made one week later. The average results, given in Table II, show an apparent displacement

TABLE II

THE AMOUNT OF DISPLACEMENT (IN MILLIMETERS) OF THE CENTER OF A STRAIGHT LINE FOLLOWING THE INSPECTION OF A CURVED LINE FOR 5 MIN

Subject	Test I	Test II
12	1.25 mm	1.25 mm
13	1.75	2
14	2	2
15	2.5	1.75
16	0	0
17	2.75	2.25
18	2	1.25
19	3	3.25
20	1.75	1.75
21	0	0
22	0	0
Average	1.545 mm	1.409 mm

slightly less than that found in the first experiment. Again the present study agrees with Gibson. However his report that the after-effect occurred in all subjects is not confirmed. Five subjects of the 22 used in this experimental series failed to report any apparent curvature of the test line.

Series II—Effect of Amount of Curvature

A new group of 12 subjects was tested for after-effect using a curved line with an 80 mm displacement at the center as a stimulus. In a second experiment they inspected a stimulus line with a displacement of 40 mm at the center. Each stimulus was exposed for 5 minutes.

The greater curvature of the first stimulus line did not increase the after-effect, but on the contrary diminished it slightly. The average after-effect of the stimulus of the greater curvature measured only 1.4 mm, while the average

after-effect of the stimulus of lesser curvature measured 1.8 mm. Concerning this point Gibson states, "A strongly curved line did not seem to give a significantly better effect than a rather weak curvature."

Series III—Duration of the Negative After-Curvature

Five experiments were carried out at weekly intervals on a new group of 9 subjects. In the first experiment the subjects repeated the tests of Series I with 5 min inspection period. In the second experiment a 30 sec interval was interposed between the inspection period and the presentation of the test line. During this interval the subject fixated a black dot on a white ground. The third experiment differed from the second in that the subject fixated the dot for 60 sec. In the fourth, the interval lasted 30 sec during which the subjects read a newspaper clipping. In the final experiment of this series the subjects read for 60 sec before the test of after-effect was made.

The results of this series are shown in Table III. The

TABLE III

CHANGE IN THE AMOUNT OF AFTER-EFFECT RESULTING FROM THE INTERPOSITION OF INTERVALS OF VARYING CHARACTER BETWEEN THE INSPECTION PERIOD AND THE TEST PERIOD

Subject	No. Interval	After 30 sec Fixation	After 60 sec Fixation	After 30 sec Reading	After 60 sec Reading
35.....	2 mm	1.25	0 mm	2	2.25
36.....	2.25	1.5	1	1.2	1.75
37.....	2.25	1.75	1.25	2	1.5
38.....	3.25	1.75	1.25	1.75	2
39.....	3	2	.75	2	1.75
40.....	2.75	1.5	1.5	1.5	1
41.....	2.25	1.5	.75	1.5	1
42.....	2.25	1.75	1	1.75	1.25
43.....		1.25	0	2	1.5
Average.....	2.5 mm	1.583 mm	.833 mm	1.777 mm	1.555 mm

average after-effect measured immediately after the 5 min inspection period was 2.5 mm; with a 30 sec fixation interval between the inspection period and the presentation of the test line the after-effect dropped to 1.58 mm; with a 60 sec

interval it dropped to .83 mm. Following the 30 and 60 sec reading intervals after-effects of 1.78 mm and 1.55 mm occurred.

The after-effect is thus proved to be strongest immediately following the inspection period and to decrease considerably within 60 sec under both conditions. It is interesting that reading periods do not decrease the after-effect as much as fixation periods of equal duration. The average after-effect is .2 mm less following the 30 sec fixation period than following the 30 sec reading period. The average after-effect following the 60 sec fixation period is .7 mm less than that following the 60 sec reading period. Although these differences between the fixation and reading periods are small, they are nevertheless, thought to be significant since every subject but one shows a greater after-effect following the 60 sec reading-period than following the 60 sec fixation period.

The present writers have no explanation for the differences in after-effect following reading and fixation. The precise relation between eye-activity and visual after-images is not known. While after-images of color and of movement generally do not appear without fixation, nevertheless a quick movement of the eyes will bring back after-images which have disappeared. A theory which attempts to explain such visual after-effects entirely in terms of eye-movement may be inadequate. The present experiment, however, shows that the character of eye-movements determines, at least in part, the nature of the after-effect.

(Manuscript received June 2, 1934)